COMPUTING CALCIUM DYNAMICS IN MAN

A thesis presented for the degree of
Doctor of Philosophy in Electrical Engineering
in the University of Canterbury,
Christchurch, New Zealand.

by

A.E. McKinnon B.E.(Hons)

1973

ABSTRACT

Current practice in the construction and use of models in medicine is considered. CAMET2 is a compartment model of calcium metabolism. Its description of the flow of material in the gut is inconsistent, and is corrected in CAMET3. In CAMET2, the rate of emptying of the stomach is proportional to the square root of its volume. In CAMET3, the distension of the duodenum inhibits the emptying of the stomach. There is experimental evidence for this mechanism. An observed 10-20 minute delay in the appearance in blood of radioactive calcium taken by mouth, is still unaccounted for.

Present methods for fitting sums of decaying exponential functions to radioactive tracer measurements are reviewed. Even when the measurement errors are as low as 2%, and the nearest exponents are in the ratio 2:1, the number of exponential components cannot be reliably determined. The parameters may be in error by 60%. The method of Lemaitre and Malengé (Lemaitre A., Malengé J-P. (1971), "An efficient method for multiexponential fitting with a computer", Cptrs. and Biomed. Res. 4: 555), is the most flexible and efficient where least squares is the criterion of goodness of fit.

Data commonly are a measure of a function which is a convolution of a desired function and a known or measured function. The Fourier transforms of both the function to which the data belong, and of the known or measured function must both be zero at the same points in the complex plane. The data are both truncated and also contain experimental error. When the error dominates the problem, the constraints on the positions of the zeros are used to improve the estimate of one

or more of the functions above. When, on the other hand, the truncation dominates, the constraints are used to estimate values outside the interval of the data; in this way the effect of the truncation is removed.

In a new version of the simulation program to solve CAMET3, the integration step length is adjusted automatically to ensure that not only is the integration efficiently performed, but also that the steps fit the inputs and outputs. A five-fold reduction in computer time results in the present case.

For the simulation of small models or parts of large models, a new interactive program SIMUL8, has been developed. The model equations are expressed simply in a FORTRAN subroutine and flexible control of graphical output on an oscilloscope display is provided.

ACKNOWLEDGEMENTS

I wish to express my gratitude to my joint supervisors, Dr R.H.T. Bates and Dr W.S. Metcalf. Their encouragement and guidance during the course of this project have been invaluable.

I am indebted to my colleagues in the Electrical Engineering Department for many stimulating discussions on aspects of the present work, particularly in relation to the computing problems which arose.

I wish to thank my wife Barbara for her patience and encouragement, and for her diligent tracing of the diagrams in this volume.

During the course of this project I was a grateful recipient of a University Grants Committee scholarship.

CONTENTS

PREFACE

The work presented in this thesis developed from consideration of problems relating to the modelling and simulation of calcium dynamics in man. It forms part of an interdisciplinary project initiated by Dr W.S. Metcalf of the Chemistry Department at the University of Canterbury. We are concerned with three aspects; the analysis of experimental data; some details of the calcium model itself; and the provision of computer facilities for solving it.

In chapter 1 we review the application of models to medical problems and we consider various aspects of model construction and use. Chapter 2 is a review of models of calcium in man, and in particular the model CAMET2, which is the basis of the interdisciplinary project. Past simulations using CAMET2 are tabulated, and its deficiencies are noted.

Tests with radioactive isotopes are common in calcium metabolism studies, and in chapter 3 we show that under certain assumptions, the amount of tracer in blood after an initial injection, is described by a sum of decaying exponential functions of time. A review of methods for solving the difficult problem of fitting exponential functions to data is given. We investigate one particular method further in chapter 4, but conclude that, although it is mathematically elegant, the inverse Laplace transform procedure suffers from severe practical limitations.

The tests in which separate oral and intravenous doses of different calcium tracers are given to patients, was the original motivation behind the work in chapters 5, 6 and 7. These chapters contain the bulk of the original material in

the present research. In the calcium test, the amounts of orally and intravenously administered tracer in the blood are related by the convolution integral. The rate at which calcium tracer enters the blood from the gut for the first time, is obtained by deconvolution. A review of current methods for this widely occurring problem concludes chapter 3.

In chapter 5 we demonstrate that there is, in principle, redundant information available in many deconvolution problems, and from this we generate consistency conditions on the given data. The consistency conditions apply to the values of the Fourier transforms of the data in the complex plane, and in chapter 5, we deviate to show that there may be some advantage in carrying out the popular Fourier method of deconvolution by inverse filtering for complex values of the argument, instead of for the usual real values. We apply our consistency conditions to X-ray diffraction data in chapter 6. Both analytically generated and actual data are used, the former successfully, the latter not. Numerous computational results are given in an attempt to arrive at a satisfactory algorithm.

We deconvolve data from the calcium absorption experiments in chapter 7, after first demonstrating our technique on analytically generated data. In this case the consistency conditions are used to compensate for the truncation, in time, of the measurement.

Chapter 8 gives details of the computational methods used for the procedures in chapters 4-7. Except as stated in chapter 8, all computer programs were written by the author, in FORTRAN for the IBM 360/44 computer.

In chapter 9 we develop the calcium model CAMET3 by modifying CAMET2. The modifications remove an inconsistency in the description of intestinal fluid flow in CAMET2, and provide a more physical description of the emptying of the stomach. The delayed appearance in blood of orally given calcium tracer is considered in the light of the model.

Chapters 10 and 11 are concerned with simulation. In chapter 10 we describe the computer program which implements CAMET3 and we demonstrate our attempts to improve its flexibility and efficiency, compared with previous versions. Interactive simulation, where possible, offers distinct advantages over simulation in 'batch' mode. In chapter 11 we describe and demonstrate an interactive simulation program which is suitable for the manipulation of small models.

Chapter 12 contains concluding remarks and suggests various areas for future research.

Oral presentations and papers on topics relevant to the material presented in this thesis are as follows:

McKinnon A.E. (1971) Mathematical models and computer simulation of biochemical dynamics. Presented at the New Zealand Endocrinology Society Conference, Christchurch, New Zealand, August 1971.

Jordan R.B., McKinnon A.E. Stomach emptying modelled on a small computer with oscilloscope display. Presented at the Quantitative Biology Meeting, Nelson, New Zealand, May 1972.

Bates R.H.T., McKinnon A.E., Napier P.J. Self-consistent deconvolution. To be submitted to Journal of Physics A (General Physics).

GLOSSARY

Unless otherwise stated abbreviations and symbols used in this thesis are as defined below. The symbols used in the models CAMET2 and CAMET3 are defined separately in Appendix I.

arg	argument, e.g. $z = \|z\| e^{\arg z}$
a_f, a_g, a_h	parameters in the Gaussian weighting functions (e.g. e^{-x^2/a_f^2}) of the Hermite function representations of f(x), g(x) and h(x).
a_i	specific activity in compartment i
A_i	coefficient of $e^{-\lambda_i t}$.
A_m	$\mathrm{Re}\{F_m\}$
B_m	$\mathrm{Im}\{F_m\}$
Br	Bromwich contour of integration; is a line parallel to the imaginary axis but to the right of all singularities of the integrand.
C, C'	contours in the complex plane
E_c	error in the corrector formula in the numerical solution of differential equations.
f(x), g(x), h(x)	functions related by the convolution integral $f(x) = \int_{-\infty}^{\infty} g(x-y)h(y)dy$.
$f(\xi)$	the Mellin transform of y(t)
$f_g(x)$	the inverse Fourier transform of $F_G(w)$
F(w), G(w), H(w)	the Fourier transforms of f(x), g(x) and h(x). e.g. $F(w) = \int_{-\infty}^{\infty} f(x)e^{j2\pi wx}dx$.
$F_\ell(s), G_\ell(s), H_\ell(s)$	the Laplace transforms of f(t), g(t), h(t) e.g. $F(s) = \int_0^{\infty} f(t)e^{-st}dt$.
F_m, G_m, H_m	coefficients in the Hermite function or Fourier series representations of f(x), g(x) and h(x).

$F_G(w)$	F(w) with M_G of its zeros replaced by the corresponding zeros of G(w)
$\mathcal{F}$	general function
$g_f(x)$	the inverse Fourier transform of $G_F(w)$
$G_F(w)$	G(w) with its zeros replaced by the corresponding zeros of F(w).
$h_0(x)$, $h_{-1}(x)$	h(x) calculated from H(u) and H(u-j) respectively.
$He_m(x)$	Hermite polynomial of order m
j	$j^2 = -1$, unless used as a subscript
$\bar{k}$	vector or matrix of unknown parameters
$K(\xi,t)$, $K^{-1}(p,\xi)$	general forward and inverse integral transform kernels
L_f, L_g, L_h	extents of f(x), g(x) and h(x).
m_f, m_g, m_h	highest order term in the Fourier series representations of f(x), g(x), h(x).
m_i	mass of material in compartment i.
M_F, M_G, M_H	the numbers of zeros in the sets ξ_F, ξ_G, ξ_H.
N	the number of exponential components
N_f, N_g, N_h	the number of samples of f(x), g(x), h(x) (e.g. $(N_f+1)T = L_f$)
p	complex time $p = t+j\tau$
$P_F(w)$	a factor of F(w) when f(x) is represented by a Fourier series
$P_{M_F}(w)$	a factor of F(w) when f(x) is represented by Hermite functions.
q_i	the amount of tracer in compartment i
$Q_F(w)$	a factor of F(w) such that $F(w) = Q_F(w)P_F(w)$.
$Q_{M_F}(x)$	a polynomial of order M_F

r_{ij}	rate of flow of material from compartment j to compartment i
$R(\zeta)$	$f(1+j\zeta)/\Gamma(1+j\zeta)$
r_{io}, R_{io}	rate of flow of material into compartment i from the environment.
RMS	root mean squared
S_{io}	rate of flow of tracer into compartment i from the environment
t	time
T	sample interval
T_Y	maximum time for which tracer data is given
u_d	width of the cosine bell window.
u_{max}	value of u at which H(u) is truncated
w	complex Fourier transform variable $w = u+jv$
$w_{F,m}, w_{G,m}, w_{H,m}$	zeros of F(w), G(w), H(w), e.g. $F(w_{F,m}) = 0$, $w_{F,m} = u_{F,m} + jv_{F,m}$.
y(t)	a solution to a model
y_n	solution to a model at the n^{th} time step when solving differential equations numerically
y^p_{n+1}	predicted solution
y^c_{n+1}	corrected solution
$Y(\alpha)$	inverse Laplace transform of y(t)
z	complex variable $z = x+jy$
β	$-\ln \alpha$
γ	$\ln t$
$\Gamma(z)$	gamma function
$\delta(t)$	Dirac delta function: $\delta(t) = 0$, $t \neq 0$; $\lim_{t\to 0} \delta(t) = \infty$; $\int_{-\infty}^{\infty} \delta(t)dt = 1$.
ζ_{max}	value of ζ at which $R(\zeta)$ is truncated

ζ_F, ζ_G, ζ_H	sets of zeros of F(w), G(w), H(w).
ζ_{FG}	set of zeros of F(w) with M_G members replaced by the corresponding zeros of G(w).
ζ_{GF}	set of zeros of F(w) which correspond to the zeros of G(w)
λ	exponent
μ_{ij}	parameter in a linear compartment model
ξ	complex Mellin transform variable $\xi = \eta+j\zeta$
ξ_F, ξ_G, ξ_H	sets of zeros of F(w), G(w), H(w) with non-zero imaginary parts
$\phi_m(t)$	arbitrary function
$\prod_{m=1}^{M}(n)$	continued product with n^{th} term omitted
*	complex conjugate, $w^* = u-jv$
ln	natural logarithm
Re{ }	real part of
Im{ }	imaginary part of
!	factorial
| |	modulus, e.g. $\|z\| = (x^2+y^2)^{\frac{1}{2}}$

C H A P T E R 1

MODELLING IN MEDICINE

Models have been used to represent systems ranging from a simple chemical reaction (Atkins, 1969, p.15), to the complex society in which we live (Watt and staff, 1970), and they form a part of many medical research programs (cf. Coleman, 1972; Garfinkel et al., 1972).

1.1 MODELLING PHILOSOPHY

There is currently no generally applied theory of models or systems. Indeed, the words "model" and "system" are themselves variously used. In this section we define them, and discuss guidelines and procedures for developing and using models.

1.1.1 Models and Systems

The word "system" has been used with reference either to reality or to a representation of it (Table 1.1). Mihram (1972) refers to the former when he defines a system as a 'collection of interdependent elements which act together in a collective effort to achieve some goal'. We use "system" with reference only to physical reality, and qualify it to focus attention on a particular aspect of reality (as, for example, in the terms "cardiovascular system", "acid-base control system" and "calcium metabolism system"). However, we acknowledge that "systems theory" (always with "systems" in the plural) is the theory of formal (mathematical) models of either real or conceptual systems (Mesarović, 1968).

Table 1.1

Use of the Word "System"

Context	Used with reference to*	Reference
General system	Rep.	Mesarović et al. (1970,p.69)
Biological system	Real	Apter (1970)
Physical system	Real	MacFarlane (1964,p.11)
Control system	Real	Elgerd (1967,p.1)
Control system	Rep.	Elgerd (1967,p.1)
Model system	Rep.	MacFarlane (1964,p.12)
Mathematical system	Rep.	MacFarlane (1964,p.12)
Linear system	Rep.	Elgerd (1967,p.45)
Hierarchical system	Rep.	Mesarović et al. (1970)
Measuring system	Real	MacFarlane (1964)
Systems theory	Rep.	Mesarović (1968)
Systems analysis	Rep.	Milsum (1970)
Systems analysis	Real	Waterman (1968)
Systems physiology	Real	Li and Urquhart (1969)
Systems synthesis	Rep.	Mihram (1972)
System identification	Real	Bekey (1970)

* Real designates reality, and Rep. designates some representation of reality.

Depending on its nature and the current state of knowledge, the system may or may not be precisely defined. In this respect Zadeh (1970) discusses 'fuzzy' systems.

Mihram (1972) gives an historical account and classification of the ways in which systems have been represented. He demonstrates that the word "model" is also variously used. Achinstein (1965) defines a model as a set of assumptions about a system, and argues that "model" and "theory" may be used interchangeably. Mesarović (1968) states that a theory is always based on an 'image' termed a model, and that any theory may be represented by a mathematical relation. He discusses this latter point in detail but does not prove it.

The definition of a model as a set of mathematical equations, is either implied or stated by authors such as Berman (1963), Li and Urquhart (1969), Birge et al. (1969) and Garfinkel (1969).

According to Apter (1970), a model functions independently of the real system, is internally consistent, and functions according to physical or mathematical laws which can be mathematically expressed. Apter's objective is to define the most useful representation of a system. Diagrams and verbal constructs, although of use, are not included among his models. He also excludes empirical equations, which are sets of symbols manipulated to fit experimental data. However, an equation which is used empirically may take the form of a physical law, and so Apter's distinctions are not altogether clear.

We define a model to be a representation in terms of mathematical equations as summarised above. We accept empiricism but aim to reduce it.

1.1.2 Types of Model

Table 1.2 lists various usages of the term "model" as we have just defined it. The subdivision of models as static, dynamic, deterministic, stochastic, continuous or discrete, has proved most useful. These terms are defined in Table 1.3 according to the characteristics of the model inputs and outputs.

Because of the various facilities available for their computer implementation, the distinction between continuous and discrete models is the most important. Continuous models may be implemented on the analogue computer (Jackson, 1960), or by using a variety of digital simulation languages (see chapter 11). For discrete models, digital simulation languages such as GPSS (Herscovitch and Schneider, 1965), GASP (Pritsker and Kiviat, 1969) and SIMSCRIPT (Kiviat et al., 1968), are widely used. Fahrland (1970) points out that discrete and continuous models are actually special cases of each other. He proposes a combined discrete and continuous simulation language, but current practice favours the languages designed specifically for one model type or the other.

We see in the following section, how the type of model developed depends not only on the system being represented, but also on the objectives and ability of the modeller.

1.1.3 Modelling Procedures and Pitfalls

Although modelling is a skill which is improved by practice (Garfinkel et al., 1972), it is profitable to analyse the process. In our view the important steps are System Analysis, Model Synthesis, Parameter Estimation, Computer Implementation, Model Testing and Prediction.

Table 1.2

Types of Model

Context	Meaning	Reference
Symbolic model	Representation by equations	Taha (1971)
Mathematical model	Physical phenomena represented by equations	Li and Urquhart (1969)
Mathematical model	Representation in terms of mathematical laws	Apter (1970)
Computer model	Equations implemented on a computer	Li and Urquhart (1969)
Simulation model	Equations implemented on a computer	Taha (1971)
Static model	Representation does not depend on time	Mihram (1972)
Dynamic model	Representation depends on time	
Deterministic model	Representation which has a unique output for a given input	
Stochastic model	Representation includes random fluctuations	
Continuous model	Representation applies at all times	
Discrete model	Representation applies only at distinct times	
Steady-state model	Representation in which quantities and flows are constant	Wigan (1972)
Boolean model	Representation in terms of Boolean functions	Schoeffler et al. (1968)
Physical model	Equations correspond well with physical phenomena	Birge et al. (1969)
Linear system (model)	Representation in terms of linear ordinary differential equations	Atkins (1969, p.19)
Simple model	Representation in terms of equations which are readily solved or easily understood	Apter (1970)
Compartmental model	Equations derived by considering the system as a set of distinct compartments	Atkins (1969) and section 3.2

Table 1.3

Definitions of Some Model Types

Model Type	Restrictions on Inputs	Properties of Outputs
Static	None	Directly related to current value of inputs.
Dynamic	None	Directly related to present and past values of inputs.
Deterministic	None	Uniquely defined by inputs.
Stochastic	None	Contain random fluctuations.
Continuous	None	Exist at all times.
Discrete	Exist only at distinct times	Exist only at distinct times.

1. System Analysis is the planning step in model development, in which the aims of the project and the scope of the model are first decided. The aims of the project may be precisely defined, as is required for an economic model to be used in making management decisions, or loosely stated, to allow the necessary flexibility when a model is being developed as part of a research project. The model CAMET2 described in section 2.2 currently accounts for some of the known facts about calcium metabolism in man. However, the aims of the project range from the satisfactory representation of all known facts about calcium metabolism, to the rapid, accurate prediction of the effects of various treatments for calcium disorders. The feasibility of any particular aim becomes apparent as the project proceeds.

The specification of model 'inputs' and 'outputs' has the effect of isolating the system being modelled from its environment, because by definition, outputs do not affect inputs. For example, fluid intake, which is an input to CAMET2, is completely unaffected by fluid loss in the urine, an output. That thirst, and hence fluid intake, is actually controlled to satisfy body fluid requirements, is not taken into account in the present model.

We further determine the scope of the model by dividing it into submodels, the details of which can be considered separately. For example, in chapter 9 we discuss the intestinal fluid dynamics of CAMET2 independently of all calcium flows and controls. McLeod (1970) refers to the isolation of submodels as the 'divide and conquer' approach.

Concurrent with the isolation of submodels, all the variables and parameters to be initially included in the

representation must be specified. While sifting the evidence for the important variables and parameters, assumptions will be made about the relative significance of various experimental results. All assumptions must be recorded for subsequent assessment.

The type of model to be developed follows naturally from the aims of the project and the assumptions made about the quantities being represented. For example, a model used in making management decisions should be stochastic so that the probability of success can be determined. On the other hand, CAMET2 is a deterministic model which has proved useful even though it represents phenomena known to have large random variations. A calcium model used for treatment studies, like an economic model, would need to account statistically for the natural variation within and between subjects.

Many of the decisions made during System Analysis are necessarily tentative. We must expect to return to the System Analysis step to reappraise the project aims and the various assumptions made in the light of subsequent experience in building the model.

2. Model Synthesis is the development of a mathematical description of the model or submodels. Brennan (1968) demonstrates synthesis either using formal mathematical equations, or an ad hoc description of the phenomenon being modelled. The latter approach directly generates a computer procedure for the solution of the model. However, we believe that it is instructive (at least for continuous models such as CAMET2) to develop the mathematical relations as an explicit set of differential equations which describe specific physical processes.

3. Parameter Estimation is required to determined the values of parameters in the mathematical description of the model, from available measurements. Parameters of submodels may be estimated separately.

The parameter estimation procedures discussed in chapter 3 (p.41) nearly all require computer implementation of the model. To increase the computational efficiency of the often repetitive solutions required, computer implementation for estimation purposes is frequently different from that intended for the fully developed model. For example, on a digital computer, the analytical solution of the model differential equations, if available, is preferred to their solution by numerical methods. In any case, the implemented model must be checked for logical errors, before any estimation is commenced. The selection of a parameter estimation algorithm is based on availability, speed, accuracy and convenience of use (chapter 3).

We may find that there is insufficient data to estimate parameter values within a given tolerance. Alternatively, the best fit of the model to the data may be judged unsatisfactory in view of the estimated experimental error. In the former case the model is 'unidentifiable' and in the latter it is 'inconsistent' (section 3.1.1). In either case we must return to the System Analysis and Model Synthesis steps and review the decisions and assumptions made.

4. Computer Implementation of the fully developed model varies according to the facilities available, the nature of the model, and the experience of the modeller. We discussed the various alternatives in section 1.1.2.

Mihram (1972) refers to the logical checking of an implemented model as 'verification'. For deterministic models tests are run for which the correct model behaviour has been determined by other means. Statistical tests such as those described by Mihram (1972) are required for checking the implementation of a stochastic model.

5. Model Testing is the comparison of model and system behaviour, with the intention of finding an experiment which the model cannot satisfactorily represent. This step is also referred to as 'Validation' (cf. Mihram, 1972; Wigan, 1972) but we avoid the term, because it suggests a search for experiments which the model can represent.

Wigan (1972) points out that frequently all available data is used during the Parameter Estimation step, so that no separate Model Testing can be done. However, this is not necessarily as bad as he suggests , because the use of all available data in the Parameter Estimation step means that the best estimate of parameter values is obtained. The important thing is the amount of independent data available compared with the number of parameters being estimated. If there is sufficient data, we can test the model as well as estimate its parameters.

When a new set of data becomes available, we can test the model either by comparing its predictions with the measurement, or by including the new information with all previous data, in a revised Parameter Estimation step. If the model is satisfactory, the latter procedure is preferred because it improves the estimate of the parameters. However, if there is a discrepancy, a careful look at the relevant parts of the model is required, remembering that the incompatibility may be caused by an earlier set of data. It could be that an effect

assumed unimportant during System Analysis, is actually highly relevant to some experimental results.

Wigan (1972) shows that the failure of a model test may be caused not only by an incorrect model, but also by using an inappropriate criterion of goodness of fit during Parameter Estimation . Unfortunately, he also reports that it is often difficult to assess the effect of the criterion of goodness of fit.

6. Prediction is the process of extracting information from a model. Frequently, predictions with biological models are used for designing further experiments on the system (Apter, 1970), so that confidence in the model is not of prime importance. This contrasts with a model used for treatment studies where much more confidence in its predictions is required.

In section 1.1.4 we describe a variety of ways for extracting information from a model.

We believe that System Analysis, Model Testing, and Prediction are the most important steps in modelling, and that faults here give rise to criticisms such as those referred to by Coleman (1972). In practice, the problems of Parameter Estimation and Computer Implementation have a short term appeal which can distract the researcher from his main objectives.

Golomb (1971) gives a clear, lighthearted summary of the pitfalls of model construction and use. In a table headed '"Dos and Don'ts of Mathematical Modelling"' (containing only "don'ts") he notes the modelling errors described more fully by others (cf. Apter, 1970; Garfinkel et al., 1972).

The latter authors discuss the important question of when work on a model should cease. If no new ideas are forthcoming about how to get further useful information out of the model (or about the system), it is pointless to continue working with it. This condition may arise because the original aims of the project have been achieved, or because further experimental data are required.

1.1.4 Model Elucidation

The tests which were performed on CAMET2 by Pearson (1972), and which are summarised in section 2.2 (p.34), are single runs of the model corresponding to various experimental conditions. There are various other analytical and computational techniques which are applied to models, and we review these here.

1. Model Transformation converts the model to a form which is mathematically equivalent, but not physically equivalent. Neer et al. (1967) express their model of radioactive calcium dynamics in man following an intravenous injection, in terms of various arrangements of four 'compartments' which contain calcium. Each arrangement is equally consistent with the experimental data used to determine it. Berman and Schoenfeld (1956) give the mathematical basis of Neer's transformations, and show that they are applicable only to a model consisting of linear differential equations (i.e. a 'linear' model).

Linear models have been subject to analysis by various well-developed techniques, especially frequency domain methods (cf. Elgerd, 1967, p.59). The transformation of a model into its frequency domain representation, is carried out in engineering mainly to assess its stability (cf. Elgerd,

1967, p.71). However, biological systems are usually highly stable and so frequency domain methods are not popular.

There is no corresponding theory applicable to the transformation of nonlinear models (i.e. models comprising nonlinear differential equations) except by first approximating them by a linear form. Again this approach is unpopular for biological systems because they are commonly highly nonlinear, and linear models representing them are frequently 'suspect' (Garfinkel et al., 1972).

2. Model Reduction is the process of decreasing the complexity of less relevant parts of a model. Chen et al. (1968) describe an algorithm for reducing the number of differentials in a linear model while retaining its principal dynamic features. The reduction of nonlinear models requires the judgement of the modeller, and any simplification must be shown not to destroy the adequacy of the model in representing the available data.

3. Sensitivity Analysis is the assessment of the dependence of a quantity of interest on each of several model parameters.

The sensitivity may be determined by repeatedly solving the model for various values of each parameter. However, even using a hybrid computer, only small problems can be dealt with, due to the excessive computation time required (Andreae et al., 1972). Alternatively, the partial derivatives of the quantity of interest with respect to each model parameter may be computed (Tomović, 1963). But because each partial derivative in general depends on all parameter values, the magnitude of the computing problem is not reduced.

When sensitivity with respect to more than two parameters is being assessed, it is difficult to display the computed information. Andreae et al. (1972) view the sensitivity of a quantity with respect to a number of parameters, by displaying it on separate graphs, one for each parameter. The parameter values are chosen randomly from within predefined ranges. Even though some information is lost, the resulting displays, which are called 'miniviews', are suggested as aids for viewing sensitivity information.

1.2 MODELS APPLIED IN MEDICINE

Table 1.3 lists a number of bodily functions for which computer-implemented models have been published. The references cited refer to recent modelling effort on each system. Models are variously used in medicine, and of the applications mentioned below, we note that the first two rely, in the main, on the ability of a model to store information (both data and hypotheses), and the last four use models predictively.

1. Teaching is aided by computer-implemented models of various bodily functions. Partridge (1972) considers that this is an efficient form of instruction, because the student can gain experience with highly complex systems, and because the speed at which information is presented to him can be matched to his ability to absorb it.
Teaching in this way is not a complete substitute for laboratory work, and there is a danger of students not realising the significance of biological variation when deterministic models are used. A number of teaching models are detailed in the journal to which the article by Partridge

Table 1.3

Published Models of Bodily Functions

System	Reference
Cardiovascular System	Boyers et al. (1972)
Respiratory System	Dionne (1972)
Body Fluid Control	Dickinson (1971)
Temperature Control	Hsu et al. (1972)
Neural Systems	Harmon and Lewis (1968)
Eye Function	Milsum (1970)
Ear Function	Geisler and Hubbard (1972)
Voice Production	Flanagan (1972)
Neuro-Muscular Control	Mains and Soechting (1971)
Intestinal Electrical Activity	Robertson-Dunn et al. (1971)
Pelvic Floor Neuromuscular Activity	Jones et al. (1971)
Postural Control	Agarwal et al. (1970)
Iron Metabolism	Sharney et al. (1971)
Calcium Metabolism	Pearson (1972)
Radionuclide Metabolism	Beach and Dyson (1971)
Thyroxine Transfer	Benetazzo et al. (1972)
Drug Distribution	Sheiner et al. (1972)
Menstrual Cycle	Bogumil et al. (1972)
Glucocorticoid System	Stokely and Howard (1972)
Gluconeogenesis	Anderson et al. (1971)
Glycosis	Garfinkel (1969)
Glood Glucose and Insulin	Ackerman et al. (1969)
Krebs Cycle	Garfinkel et al. (1969)
Growth Hormone Regulation	Howard and Young (1970)
Protein Metabolism	Janes and Carson (1971)

(1972) is an introduction.

2. A model parameter is a useful Diagnostic Index, if its value gives a reliable indication of the disease state of a patient. The value of the parameter is estimated from a measurement on the patient, and it is then compared with previous estimates of the same parameter for both healthy and diseased subjects. Essentially one asks how the value of the parameter affects the probability of each possible diagnosis. If it markedly affects the probability of one or more diagnoses, it is a useful diagnostic tool.

The use of model parameters as diagnostic indices is related to the whole question of the application of computers to medical diagnosis. This problem is currently receiving much attention (cf. Gleser and Collen, 1972; Warner et al., 1972).

3. Hypothesis Testing using a model may occur as part of a research program. The model is used to predict the consequences of hypotheses about the operation of the system it represents, and experimental effort is aimed at showing the predictions to be in error, thus refuting the hypothesis. Although researchers rarely attempt to estimate the degree to which modelling has assisted them, Coleman (1972) discusses the general benefits which accrue, and cites some examples where he considers the approach has been successful.

4. Device Testing of an artificial replacement for a defective part of the body is sometimes carried out using a model of the system with which the device must interact. In this case, adequate models of both the artificial aid and the relevant bodily function are required. Spyker (1970) describes work with a cardiovascular model to assess

circulatory 'assist' devices. Abbrecht and Prodany (1971) use a model to assess the performance of an artifical kidney in man.

5. Deduction from Measurement is the use of models to infer quantities that are not easily measured, from those that are. An example is the use of the two radioactive isotope measurements to obtain an assessment of the rate of absorption of calcium from the gut, as described in section 7.4. Jordan (1973) uses a model to determine the concentration of biologically active adrenocorticotrophic hormone (ACTH) in plasma, from a radioimmunoassay sensitive to both ACTH itself and to a fragment of it. Such models can be tested by direct measurement of the derived quantity, if this is possible. Unfortunately, the error in the measurements frequently appears magnified in the derived quantity (as in, for example, section 7.4), so that direct measurement is preferred if it is practical.

6. Treatment Studies with a model provide a means for assessing the effect of proposed medication. Of the various models listed in Table 1.3 (p.15), those of drug distribution are most widely used for treatment studies. The subject of pharmacokinetics (cf. Wagner , 1968) is concerned with the 'dynamics of drug distribution in man and the formulation of models to describe the processes which occur'. Pharmacokinetic models are used to determine the dosage required to achieve a particular drug concentration at a specified organ, but are usually not used to predict pharmacological effects.

Where the effect of treatment is not expected to be seen for some years, a rapid artificial assessment by computer may be a better guide than is observation of the patient.

CHAPTER 2

MODELS OF CALCIUM METABOLISM

Calcium plays an important role in many cellular processes, and is a primary constituent of the skeleton. The concentration of calcium in plasma, intracellular fluid, and extracellular fluid stays remarkably constant throughout the lifetime of a human being, and indeed must do so for the correct operation of cellular mechanisms (cf. Bianchi, 1968). Disorders of calcium metabolism are not always associated with an abnormal plasma calcium concentration. Renal stones (cf. Vaughan, 1970, p.109) and osteoporosis (cf. Vaughan, 1970, p.224) are ailments in this category. Osteoporosis occurs mostly in older people and is both difficult to detect in its early stages, and difficult to treat (cf. Vaughan, 1970, p.217). Computer models provide a convenient basis on which to develop an understanding of the complex interactions involved in calcium metabolism, from the many, and sometimes inconsistent, experimental results reported in the literature.

2.1 OTHER CALCIUM MODELS

The disappearance of radioactive Ca^{47} from plasma following an intravenous injection, is a common measurement of the dynamics of calcium in man, and the results obtained are described by a number of published models. Of these, models consisting of several 'compartments' (see section 3.2) are the most popular (Table 2.1), but noncompartmental models consisting of power functions have also been used. Both the compartmental and noncompartmental models are highly empirical, serving only to represent the result of the tracer

Table 2.1

Models of Calcium Metabolism

Plasma Compartmental Models

Bauer et al. (1957)

Aubert and Milhaud (1960)

Bronner and Aubert (1965)

Cohn et al. (1965)

Neer et al. (1967)

Massin et al. (1968)

Phang et al. (1969)

Gonick and Brown (1970)

Ramberg et al. (1970)

Intestinal Compartmental Models

Birge et al. (1969)

Marshall and Nordin (1969)

Regulatory Models

Roston (1959)

Copp et al. (1960)

Aubert and Bronner (1965)

Riggs (1966)

Noncompartmental Models

Marshall (1964)

Ackerman et al. (1967)

Anderson et al. (1967)

Marshall and Onkelinx (1968)

Burkinshaw et al. (1969)

test. The reviews by Heaney (1963), Livesey (1970) and Vaughan (1970, p.139) show that attempts to find a satisfactory physical basis for the mathematically determined compartments have failed, and that the physical significance of derived quantities such as 'bone accretion' is not clear. When tracer is injected into the blood, it quickly distributes itself throughout a volume which is unknown, but is larger than the vascular space. Therefore, the total amount of tracer in the first compartment at any time cannot be measured; specific activity (the ratio of tracer to stable isotope) is the only information available. We see in section 3.2.1 (p. 47) that calcium flow rates between compartments can be estimated from a specific activity measurement, only if they are effectively constant throughout the experiment. This assumption is commonly made when interpreting tracer kinetic data. Therefore, we can infer little about the mechanisms which control the plasma concentration of calcium, from typical tracer measurements.

Regulatory models, on the other hand, describe to some degree the observed behaviour when the plasma concentration of calcium is disturbed from its normal level. They attempt to represent the various interacting mechanisms; their data coming from tests on humans (where possible), and on laboratory animals.

The basis and performance of most of the models listed in Table 2.1 are reviewed in detail by Livesey (1970). However, two further regulatory models have been published since (Wolaver, 1972, and Powell, 1972). As in the regulatory models of Table 2.1, the controlled quantity is the plasma (ionic) calcium concentration. Its increase above normal

causes secretion of the hormone calcitonin (CT), and its decrease causes the secretion of parathyroid hormone (PTH). Wolaver (1972) represents the distribution of calcium, PTH and CT in two compartments, intravascular and extravascular. Powell (1972) represents them only in a plasma compartment. Both authors include the effects of PTH and CT in increasing and decreasing, respectively, the nett transfer of calcium from bone to plasma. The lowering of the renal excretion of calcium in the presence of PTH is also included in their models. Powell (1972) goes further, and in his initial formulation includes an increased intestinal absorption of calcium in the presence of PTH, and a lowering of both urinary calcium and calcium loss to the intestine in the presence of CT. Unfortunately, he finds it necessary to combine a number of his parameters together, which means that the detailed effect of PTH and/or CT at various sites (such as the kidney) can no longer be deduced from his model. He does not complete the estimation of the model parameters but indicates how this can be done from existing data and data which could be obtained from experiments on laboratory animals using current techniques.

The article by Wolaver (1972) is an abstract to a conference proceedings and so details of some mathematical expressions (such as urinary calcium excretion as a function of PTH and plasma calcium), and parameter values are not given. Neither Wolaver (1972) nor Powell (1972) include the dynamics of intestinal calcium in their models. Powell assumes that a certain fraction (subject to PTH control) of ingested calcium is absorbed, whereas Wolaver has calcium absorption as a model input.

2.2 CAMET2

The present model of calcium metabolism is part of an on-going project initiated by Dr W.S. Metcalf of the Chemistry Department of the University of Canterbury, and is the subject of two research dissertations (Livesey, 1970 and Pearson, 1972). Livesey developed the model CAMET, which was subsequently improved and extended as CAMET2 by Pearson.

Both CAMET and CAMET2 combine the various types of model given in Table 2.1 (except the power function models). Although, Livesey (1970) included hormone action on bone, kidney and gut absorption, in the formulation of CAMET, the lack of suitable data at that time prevented the quantification of these effects. The major contributions to the model by Pearson (1972) are the inclusion of the quantitative details of hormone (i.e. PTH and CT) action on the renal excretion of calcium, and the imposition of controlled bone deposition rates according to the availability of sites in the collagen matrix.

CAMET2 consists of 12 compartments containing calcium (Fig. 2.2). The first 11 of these are true homogeneous compartments as defined in section 3.2. (p.44) The twelfth is an 'unmixed compartment' such that during the period of a tracer study (Neer et al., 1967, collected data over 20 days) the amount of tracer leaving it is negligible compared with that entering. The existence of an unmixed compartment is the basis of the definition of bone accretion (cf. Heaney, 1963). The compartments numbered 1-4 are empirically derived from tracer data. Compartment 1 certainly includes the calcium in the vascular space, but its actual physical counterpart (if any) and that of compartments 2, 3 and 4 is uncertain.

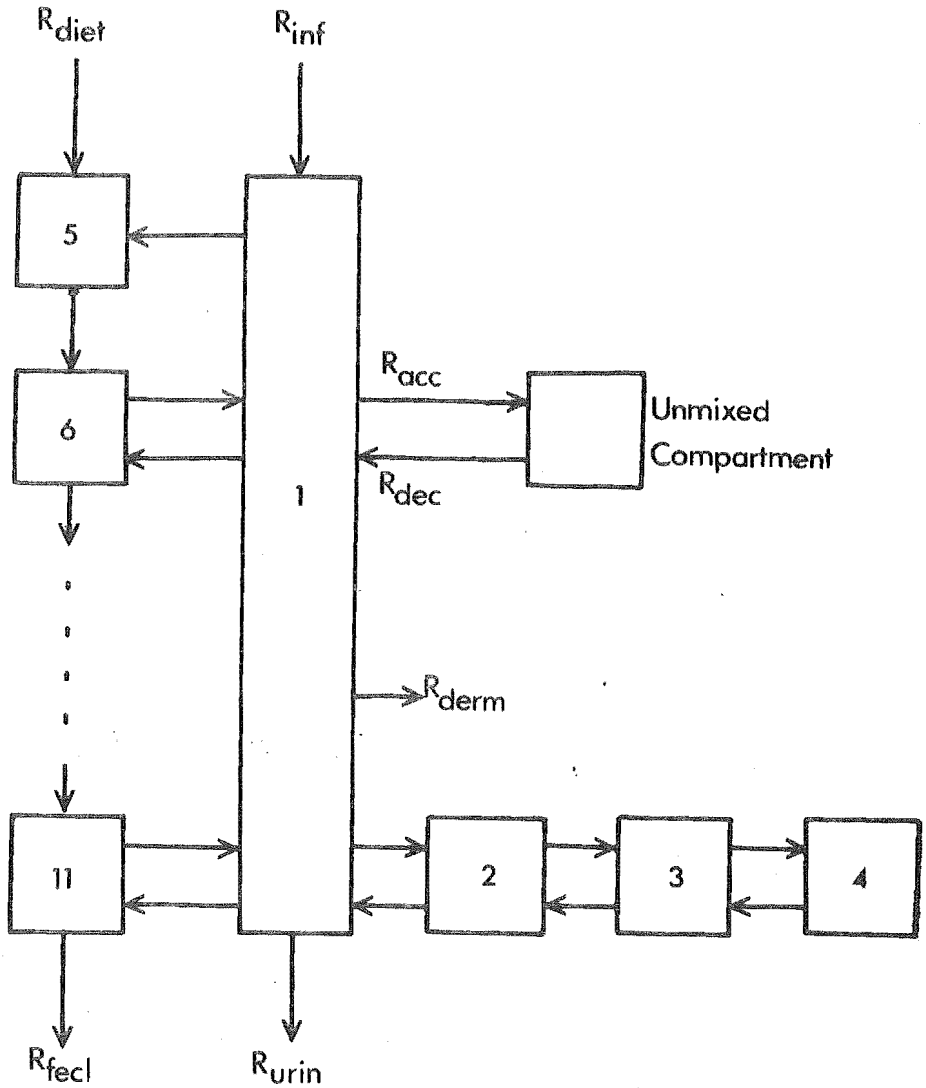


Figure 2.2 Calcium compartments and flows in CAMET2

Pearson (1972) deleted compartments 3 and 4 from the model describing the dynamics of stable calcium, but they remain in CAMET2 to describe the dynamics of tracer, because ion exchange affects the latter, but not the former. Compartment 5 represents calcium in the stomach and compartments 6-11 represent calcium in equal length sections of the intestine from the upper duodenum to the distal ileum. Corresponding to each calcium compartment 5-11 is a compartment of fluid,

so that the dynamics of material flow in the intestine are described.

Table 2.3 lists the equations which describe CAMET2. The symbols used are defined in Appendix I, but some of the important ones are:

M_I the quantity of calcium in compartment I

MT_I the quantity of tracer in compartment I

$R_{I,J}$ the rate of transfer of calcium from compartment I to compartment J

$RT_{I,J}$ the rate of transfer of tracer from compartment I to compartment J

$RF_{I,J}$ the rate of transfer of fluid from compartment I to compartment J

$\dot{x}$ the first derivative of x with respect to time.

The equations in Table 2.3 denoted with an asterisk are the same as those in CAMET (Livesey, 1970). Where applicable, quantities are normalised to unit body weight. A body weight of 65 kg is used if the actual body weight is not known.

Table 2.3

CAMET2 Equations

Differential Equations

Plasma and exchanging calcium

$$\dot{M}_1 = R_{2,1} + R_{inf} + R_{dec} + R_{abs} - R_{1,2} - R_{acc} - R_{gsec} - R_{urin} - R_{derm} \quad (2.1)$$

$$\dot{MT}_1 = RT_{2,1} + RT_{inf} + RT_{abs} - RT_{1,2} - RT_{acc} - RT_{gsec} - RT_{urin} - RT_{derm} \quad (2.2)$$

$$\dot{M}_2 = R_{1,2} - R_{2,1} \quad (2.3)$$

$$*\dot{MT}_2 = RT_{1,2} + RT_{3,2} - RT_{2,1} - RT_{2,3} \quad (2.4)$$

$$*\dot{MT}_3 = RT_{2,3} + RT_{4,3} - RT_{3,2} - RT_{3,4} \quad (2.5)$$

$$*\dot{MT}_4 = RT_{3,4} - RT_{4,3} \quad (2.6)$$

Stomach

$$*\dot{M}_5 = R_{diet} + R_{1,5} - R_{5,6} \quad (2.7)$$

$$*\dot{MT}_5 = RT_{diet} + RT_{1,5} - RT_{5,6} \quad (2.8)$$

$$*\dot{V}_5 = RF_{diet} + RF_{1,5} - RF_{5,6} \quad (2.9)$$

Intestine

$$*\dot{M}_I = R_{(I-1),I} + R_{1,I} - R_{I,(I+1)} - R_{I,1} \quad (2.10)$$

$$6 \leqslant I \leqslant 10$$

$$*\dot{MT}_I = RT_{(I-1),I} + RT_{1,I} - RT_{I,(I+1)} - RT_{I,1} \quad (2.11)$$

$$6 \leqslant I \leqslant 10$$

$$*\dot{M}_{11} = R_{10,11} + R_{1,11} - R_{fecl} - R_{11,1} \quad (2.12)$$

$$*MT_{11} = RT_{10,11} + RT_{1,11} - RT_{fecl} - RT_{11,1} \quad (2.13)$$

Hormones

$$\dot{M}_c = R_{sc} + R_{cinf} - R_{dc} \quad (2.14)$$

$$\dot{M}_p = R_{sp} + R_{pinf} - R_{dp} \quad (2.15)$$

Bone

$$\dot{M}_{bca} = R_{acc} - R_{dec} \quad (2.16)$$

$$\dot{M}_{bcoll} = R_{coacc} - R_{codec} \quad (2.17)$$

Rate Expressions and Volumes

Plasma and Exchanging Calcium

$$*R_{1,2} = M_1 K_{1,2} \quad (2.18)$$

$$*R_{2,1} = M_2 K_{2,1} \quad (2.19)$$

$$*RT_{I,J} = MT_I K_{I,J} \quad 1 \leqslant I \leqslant 4,\ 1 \leqslant J \leqslant 4,\ I = J \pm 1 \quad (2.20)$$

$$*R_{gsec} = \sum_{I=5}^{11} R_{1,I} \quad (2.21)$$

$$*RT_{gsec} = \sum_{I=5}^{11} RT_{1,I} \quad (2.22)$$

Stomach

$$*RF_{5,6} = FV_5^{\frac{1}{2}} \quad (2.23)$$

$$*R_{5,6} = RF_{5,6} C_5 \quad (2.24)$$

$$*RT_{5,6} = R_{5,6} SA_5 \quad (2.25)$$

$$R_{1,5} = K_c RG_5 \quad (2.26)$$

$$RT_{1,5} = R_{1,5} SA_1 \quad (2.27)$$

Intestine

$$RF_{1,6} = 3RF_{5,6} + G_1 \quad (2.28)$$

$$R_{1,6} = K_c(G_2 RF_{1,6} + G_3) \quad (2.29)$$

$$R_{1,I} = K_c RG_I \qquad 7 \leqslant I \leqslant 11 \quad (2.30)$$

$$^*RF_{I,(I+1)} = RF_{(I-1),I}(1-AF_I) + RF_{1,I} \qquad 6 \leqslant I \leqslant 10 \quad (2.31)$$

$$^*RF_{fecl} = RF_{10,11}(1-AF_{11}) + RF_{1,11} \quad (2.32)$$

$$^*V_I = B_{I,1}RF_{I,(I+1)}/(B_{I,2} + RF_{I,(I+1)}) \qquad 6 \leqslant I \leqslant 10 \quad (2.33)$$

$$^*V_{11} = B_{11,1}RF_{fecl}/(B_{11,2} + RF_{fecl}) \quad (2.34)$$

$$^*R_{I,(I+1)} = RF_{I,(I+1)}C_I \qquad 6 \leqslant I \leqslant 10 \quad (2.35)$$

$$^*RT_{I,(I+1)} = R_{I,(I+1)}SA_I \qquad 6 \leqslant I \leqslant 10 \quad (2.36)$$

$$^*R_{fecl} = RF_{fecl}C_{11} \quad (2.37)$$

$$^*RT_{fecl} = R_{fecl}SA_{11} \quad (2.38)$$

$$^*R_{I,1} = E_{I,1}C_I/(E_{I,2} + C_I) \qquad 6 \leqslant I \leqslant 11 \quad (2.39)$$

$$^*RT_{I,1} = R_{I,1}SA_I \qquad 6 \leqslant I \leqslant 11 \quad (2.40)$$

$$RT_{1,I} = R_{1,I}SA_1 \qquad 5 \leqslant I \leqslant 11 \quad (2.41)$$

$$R_{abs} = \sum_{I=6}^{11} R_{I,1} \quad (2.42)$$

$$RT_{abs} = \sum_{I=6}^{11} RT_{I,1} \quad (2.43)$$

RG_I and RF_I are constant rates - see table 2.4.

Hormones

$$R_{sc} = 0.0 \qquad C_1 \leq A_{22} \tag{2.44}$$

$$R_{sc} = A_{23}(C_1 - A_{22}) \qquad C_1 > A_{22} \tag{2.45}$$

$$R_{dc} = K_{dc}M_c \tag{2.46}$$

$$R_{sp} = A_{25}(A_{24} - C_1) \qquad C_1 < A_{24} \tag{2.47}$$

$$R_{sp} = 0.0 \qquad C_1 \geq A_{24} \tag{2.48}$$

$$R_{dp} = K_{dp}M_p \tag{2.49}$$

Bone

$$R_{coacc} = A_4 \tag{2.50}$$

$$R_{codec} = A_1/K_c \tag{2.51}$$

$$R_{acc} = K_c K_{oss} U_{mcoll} \tag{2.52}$$

$$*RT_{acc} = K_{acc} C_1 \quad SA_1 \tag{2.53}$$

$$R_{dec} = K_{rp} R_{codec} \tag{2.54}$$

Urinary Excretion of Calcium

$$R_{urin} = (A_6 C_1 - A_7)/((1 + A_8 C_p)(1 - A_9 C_c)) \tag{2.55}$$

$$*RT_{urin} = R_{urin} SA_1 \tag{2.56}$$

Dermal Excretion of Calcium

$$R_{derm} = K_d C_1 \tag{2.57}$$

$$RT_{derm} = R_{derm} SA_1 \tag{2.58}$$

Concentrations and Specific Activities

$$^{*}C_1 = M_1/V_1 \tag{2.59}$$

$$^{*}C_I = M_I/V_I \qquad 5 < I < 11 \tag{2.60}$$

$$C_c = M_c/V_c \tag{2.61}$$

$$C_p = M_p/V_p \tag{2.62}$$

$$K_c = C_1/\bar{C}_1 \text{ where } \bar{C}_1 = 2.5 \text{ m.moles/litre} \tag{2.63}$$

$$^{*}SA_I = MT_I/M_I \qquad 1 < I < 11 \tag{2.64}$$

Quantity of Unmineralized Bone Collagen

$$U_{mcoll} = M_{bcoll} - K_{pr}M_{bca} \tag{2.65}$$

Total Tracer Calcium

$$MT_{tot} = \sum_{I=1}^{11} MT_I \tag{2.66}$$

Calcium Balance

$$C_{bal} = S_{diet} + S_{inf} - S_{fecl} - S_{urin} - S_{derm} \tag{2.67}$$

$$CT_{bal} = ST_{diet} + ST_{inf} + MT_{tot} - ST_{fecl} - ST_{urin} - ST_{derm} \tag{2.68}$$

The basic set of model parameters and initial conditions for CAMET2 is given in Table 2.4. Where a parameter is the same as used in CAMET it is marked with an asterisk. The initial conditions for stomach fluid volume and the calcium in compartments 6-11, are the 'steady state' values obtained when the model is run for a long time with no inputs.

Pearson (1972) reports the details of 15 simulations which he conducted using CAMET2. The nature of the test on the model, and the conclusions he draws are summarised in Table 2.5, where we note with an asterisk any simulations similar to those previously conducted using CAMET (Livesey, 1970). Some of the simulations listed in Table 2.5 serve more to highlight a phenomenon already suspected from the literature, than to provide new information.

CAMET2 represents a significant advance over CAMET in that it now adequately describes a wider range of measured results. Pearson (1972) suggests various improvements such as; the inclusion of CT and PTH action on bone; the introduction of vitamin D and its action on intestinal absorption; the inclusion of a variable plasma volume, to account for water movement into and out of the intestine, and to improve the kidney model.

He points out two deficiencies in the gastro-intestinal section of the model, which we consider further in chapter 9. The first arises because of an inconsistency in the equations describing intestinal fluid flow. We note from equation (2.31) that the outflow ($RF_{I,(I+1)}$) of fluid from an intestinal compartment is equal to the inflow ($RF_{(I-1),I} + RF_{1,I}$) minus what is absorbed ($AF_I RF_{(I-1),I}$).

Table 2.4

CAMET2 Model Parameters and Initial Conditions

Parameter	Value	Units
Differential Equation Initial Conditions		
$*M_1$	0.45	m.moles/kg
$*M_2$	0.52	"
$*M_3$	1.23	"
$*M_4$	0.87	"
M_5	0.28×10^{-3}	"
M_6	0.11×10^{-2}	"
M_7	0.65×10^{-3}	"
M_8	0.21×10^{-3}	"
M_9	0.10×10^{-3}	"
M_{10}	0.64×10^{-4}	"
M_{11}	0.40×10^{-4}	"
$*MT_1$ to MT_{11}	0.0	"
M_c	7.1	ng/kg
M_p	93.0	ng/kg
M_{bca}	0.4242	moles/kg
M_{bcoll}	21.23	g/kg
V_5	0.187×10^{-3}	litres/kg
Rate Constants		
$*K_{1,2}$	28.9	day^{-1}
$*K_{2,1}$	25.2	"
$*K_{2,3}$	3.10	"
$*K_{3,2}$	1.30	"
$*K_{3,4}$	0.127	"

Parameter	Value	Units
Rate Constants (contd)		
$*K_{4,3}$	0.180	day^{-1}
$*K_{acc}$	0.078	"
K_d	0.0153	litres/kg/day
K_{dc}	41.7	day^{-1}
K_{dp}	36.1	"
K_{oss}	0.60×10^{-2}	litres/kg/day
Intestinal Fluid Secretion Rates		
$RF_{1,5}$	0.041	litres/kg/day
$RF_{1,7}$	0.011	"
$RF_{1,8}$ to $RF_{1,11}$	0.0024	"
Intestinal Calcium Secretion Rates		
RG_5	0.062	m.moles/kg/day
RG_7	0.016	"
RG_8	0.0036	"
RG_9 to RG_{11}	0.0036	"
Duodenal Secretion Parameters		
G_1	-0.078	litres/kg/day
G_2	0.83	m.moles/litre
G_3	0.013	m.moles/kg/day
Fluid Flow Constants		
AF_6 to AF_{11}	0.48	-
$*B_{6,1}$ to $B_{8,1}$	0.004	litres/kg
$*B_{9,1}$ to $B_{11,1}$	0.0025	"
$*B_{6,2}$ to $B_{8,2}$	0.14	litres/kg/day
$*B_{9,2}$ to $B_{11,2}$	0.12	"
F	3.0	$litres^{\frac{1}{2}}/kg^{\frac{1}{2}}/day$

Parameter	Value	Units
Hormone Effect Constants		
A_1	0.58×10^{-2}	g/kg/day
A_4	0.58×10^{-2}	"
A_6	0.262	litres/kg/day
A_7	0.405	m.moles/kg/day
A_8	3.84	mℓ/ng
A_9	0.847	"
A_{22}	1.7	m.moles/litre
A_{23}	0.37×10^{3}	ng litre m.mole^{-1} kg^{-1} day^{-1}
A_{24}	3.0	m.moles/litre
A_{25}	0.67×10^{4}	ng litre m.mole^{-1} kg^{-1} day^{-1}
Intestinal Absorption Parameters		
$*E_{6,1}$ to $E_{8,1}$	0.19	m.moles/kg/day
$*E_{9,1}$ to $E_{11,1}$	0.10	"
$*E_{6,2}$ to $E_{8,2}$	2.2	m.moles/litre
$*E_{9,2}$ to $E_{11,2}$	5.0	"
Volumes		
$*V_1$	0.18	litres/kg
V_c	71.0	ml/kg
V_p	155.0	"
Calcium/Collagen Bone Ratios		
K_{pr}	50.0	g/mole
K_{rp}	0.02	mole/g

Table 2.5

Summary of the Simulations Using CAMET2 reported by Pearson (1972)

Test on CAMET2	References	Results and Conclusions
1. 4 hr infusion of calcium	Nordin & Fraser (1954); Nordin & Smith (1965); Arnaud et al. (1971)	The initial calcium excretion predicted by CAMET2 is a little low. Increase in urinary calcium is the main response to calcium overload in man.
2. As for 1 but for thyroidectomised case (i.e. no CT)	Ibbertson et al. (1967); O'Brien & McIntosh (1967)	CT action on kidney during high blood calcium is probably underestimated in CAMET2.
3. Parathyroidectomy (i.e. no PTH)	Canary et al. (1962) Biddulph et al. (1970)	PTH action on the kidney is more important than on bone, especially in the short term.
4. 2 hr infusion of the calcium complexing agent ethylene-diamine tetraacetic acid (EDTA)	Copp (1960) Jones & Fourman (1963)	Plasma calcium in CAMET2 fails to return to normal after the infusion because no bone action of PTH is included in the model.
5. As for 4 but with parathyroidectomy (i.e. no PTH)	Copp (1960)	The agreement between model predictions and experiment supports the conclusion of simulation 4.
6. 3 meal day with 2 separate calcium infusions. Intestinal secretion rates of calcium are made to vary with the plasma concentration	Baylor et al. (1950)	It is not clear whether intestinal secretion rates of calcium should depend on the plasma concentration.
7. 3 meal day but with calcium tracer in plasma forced (artificially) to remain constant.	Heaney and Skillman (1964)	Standard clinical tests for the calcium which leaves plasma and appears in the faeces are not as accurate as is often assumed. Current methods for estimating intestinal secretions of calcium are also only approximate.
8* As for 7 except that plasma tracer is allowed to fall naturally.	Heaney and Skillman (1964)	
9. 24 hr fast.		CAMET2 predicts falls in both blood and urinary calcium. Some experimental evidence suggests that urinary calcium ultimately returns to about its normal level.

Test on CAMET2	References	Results and Conclusions
10* 3 daily meals, low in calcium, for 2 days.	MacFadyen et al. (1965); Nordin & Smith (1965)	Experiments with animals and humans on low calcium diets should provide useful insight into calcium metabolism.
11. Three meal day.		CAMET2 predictions compare satisfactorily with various experimental results, except that intestinal absorption is slightly high.
12* Single oral dose of tracer with some stable calcium.	Jaworski et al. (1963)	Ingestion of calcium with or soon after an oral tracer dose, significantly reduces the amount of tracer adsorbed into the blood.
13. Single oral dose of tracer observed for 3 days of 3 meals each.	Jaworski et al. (1963)	
14. A single oral tracer dose with calcium load, but without further meals for 24 hrs.	Mautalen et al. (1969)	Predicted plasma specific activity compares well with measurement.
15* Intravenous injection of calcium tracer followed for 3 days of 3 meals each.	Neer et al. (1967)	CAMET2 satisfactorily predicts the results on which CAMET is based.

Mass balance requires that the volume V_I therefore remains constant. However, equations (2.33) and (2.34) calculate the compartment volumes V_I from the varying flows $RF_{I,(I+1)}$. The volumes are therefore no longer constant, which means that equations (2.31), (2.32), (2.33) and (2.34) are physically inconsistent. Pearson (1972) notes some anomalies in the behaviour of CAMET2 due to the inconsistent equations. We correct them in chapter 9.

The second aspect of CAMET2 discussed in chapter 9 relates to its inability to correctly predict a 15-20 min. delay in the appearance of ingested calcium tracer in plasma. The delay is well documented (cf. Caniggia et al., 1963; Birge et al., 1969), but the reason for it is not clear.

CHAPTER 3

FITTING MODELS TO DATA

A general expression for a model consisting of ordinary differential equations (we do not consider those involving partial differential equations) is given by

$$\frac{d\bar{x}(t)}{dt} = \bar{\mathcal{F}}\,(\bar{x}(t),\ \bar{r}(t),\ \bar{k},\ t) \qquad (3.1)$$

and

$$\bar{y}(t) = \bar{\mathcal{H}}(\bar{x}(t),\ \bar{r}(t)). \qquad (3.2)$$

In equations (3.1) and (3.2) $\bar{x}$ is defined as the vector (dimension N) of 'state variables' (Elgerd, 1967, p.33), $\bar{r}$ is the vector (dimension L) of time-varying inputs, $\bar{y}$ is the vector (dimension J) of model outputs, and $\bar{k}$ is the vector (dimension M) of model parameters. $\bar{\mathcal{F}}$ and $\bar{\mathcal{H}}$ denote functional dependency, and t represents time. The number of state variables N is defined as the order of the model.

Model construction is the process of determining $\bar{\mathcal{F}}$ and $\bar{\mathcal{H}}$, and the parameters $\bar{k}$. Currently, there is no known way of directly determining $\bar{\mathcal{F}}$ and $\bar{\mathcal{H}}$ from measured data, except when the range of possible forms for $\bar{\mathcal{F}}$ and $\bar{\mathcal{H}}$ is severely restricted. The usual approach is to postulate $\bar{\mathcal{F}}$ and $\bar{\mathcal{H}}$ during the Model Synthesis step (section 1.1.3 (p.8)), and subsequently test the hypothesis during Parameter Estimation and Model Testing. When a postulate for $\bar{\mathcal{F}}$ and $\bar{\mathcal{H}}$ is available, it is necessary to estimate the parameters $\bar{k}$ so that, in some sense, the model best 'fits' the data. In this chapter we are mainly concerned with the latter problem which is variously referred to as 'system identification' (cf. Bekey, 1970), 'parameter estimation' (cf. Eykhoff, 1963), and 'parameter

identification' (cf. Burrus et al., 1971).

3.1 GENERAL CONSIDERATIONS

3.1.1 Identifiability, Uniqueness and Consistency

Suppose we have the model

$$\frac{dy_1}{dt} = -(k_1 + k_2)y_1 \tag{3.3}$$

$$\frac{dy_2}{dt} = k_1y_1 - k_3y_2 \tag{3.4}$$

which, with suitable values of k_1, k_2 and k_3, we know to be an accurate representation of a hypothetical 'system'. We perform an 'experiment' on the system by setting

$$y_1(0) = 0 \quad \text{and} \quad y_2(0) = 1, \tag{3.5}$$

and measuring $y_1(t)$ and $y_2(t)$. We see by inspection of (3.5), (3.3) and (3.4), that $\frac{dy_1}{dt}$ and $y_1(t)$ are always zero, and that $y_2(t)$ is independent of both k_1 and k_2. Therefore, estimates of k_1 and k_2 cannot be obtained from the measurements and so the system is 'unidentifiable' under the given conditions. Lee (1964, p.83) shows that a system is identifiable if the initial values excite all of its 'modes', and he gives appropriate mathematical conditions. His formulae apply only to models expressed by linear differential equations (as in our example), there being no equivalent formulae for the general model (3.1) and (3.2).

To illustrate uniqueness, we consider the model

$$\frac{dy_1}{dt} = k_2y_2 - k_3y_1 \tag{3.6}$$

$$\frac{dy_2}{dt} = k_1y_1 - k_4y_2 \tag{3.7}$$

with the initial values

$$y_1(0) = 1 \quad \text{and} \quad y_2(0) = 0. \tag{3.8}$$

Using standard procedures for solving linear differential equations (Elgerd, 1967, p.60) we find

$$y_1 = A_1 e^{-\lambda_1 t} + A_2 e^{-\lambda_2 t} \tag{3.9}$$

where

$$\lambda_1 = \tfrac{1}{2}(k_3+k_4) - \tfrac{1}{2}\sqrt{(k_3+k_4)^2 - 4k_3k_4 + 4k_2k_1}, \tag{3.10}$$

$$\lambda_2 = \tfrac{1}{2}(k_3+k_4) + \tfrac{1}{2}\sqrt{(k_3+k_4)^2 - 4k_3k_4 + 4k_2k_1}, \tag{3.11}$$

$$A_1 = \frac{k_3 - \lambda_2}{\lambda_1 - \lambda_2}, \tag{3.12}$$

and

$$A_2 = \frac{k_3 - \lambda_1}{\lambda_2 - \lambda_1}, \tag{3.13}$$

provided k_1, k_2, k_3 and k_4 are such that

$$\lambda_1 \neq \lambda_2. \tag{3.14}$$

We note that

$$A_1 + A_2 = y_1(0) = 1. \tag{3.15}$$

Therefore only three independent parameters (λ_1, λ_2 and A_1 or A_2) can be determined from any number of measurements of $y_1(t)$, providing insufficient information to uniquely define the four unknown constants k_1, k_2, k_3 and k_4.

Although our example consists of linear differential equations, the problem of 'structural identifiability' (Bellman and Aström, 1970) which it illustrates, occurs also in the general model (3.1) and (3.2). An analytical solution

to the model is always required to find out if the parameters can be determined uniquely. If not, uniqueness is obtained by preassigning values to some parameters (Bellman and Aström, 1970; Berman, 1963). Experimentation with the assigned values may be required, because a careless choice can make it impossible for the model to represent the data. (Berman, 1963). If in our example, we can state that

$$k_1, k_2, k_3, k_4 \geq 0 \qquad (3.16)$$

(these physical constraints apply for linear compartment models; see section 3.2.1, p.50) the parameter space in which k_1, k_2, k_3, k_4 lie is determined by the matrix transformations used by Berman and Schoenfeld (1956).

We have shown that the parameters k_1, k_2, k_3, and k_4 of our example (3.6) and (3.7) cannot be determined uniquely from a measurement of $y_1(t)$ alone, irrespective of its accuracy. Nonuniqueness also occurs if, due to error, there is insufficient information in the data. Berman (1963) shows that parameter estimation is fraught with computational difficulties which are resolved by preassigning either a value, or a precision, to some parameters. However, we note that if the nonuniqueness is of the type which occurs even with perfect data, a preassignment of precision will not improve the computations.

A model is inconsistent if it has insufficient freedom to allow it to adjust to the data. The condition is evidenced by the presence of systematic deviations between the best model prediction and the data (Berman, 1963). It is rectified by redesigning the model to allow more degrees of freedom.

3.1.2 Fitting Nonlinear Models

The basis of methods for estimating $\bar{k}$ in (3.1) and (3.2) (p.37) is shown in Fig. 3.1.

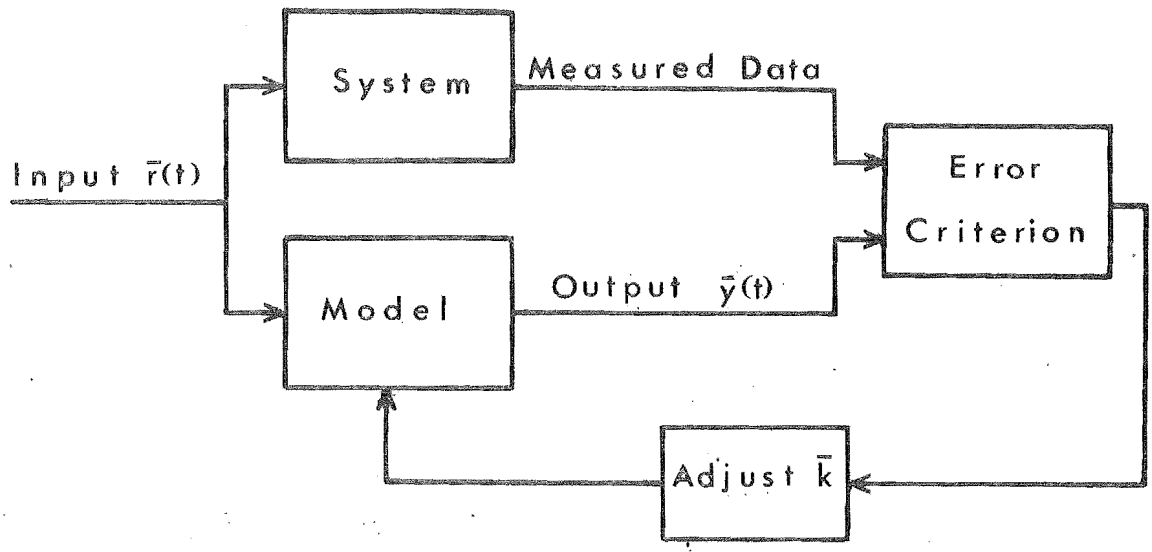


Fig. 3.1 Basic procedure for parameter estimation.

The $\bar{k}$ vector is iteratively adjusted between successive solutions of the model to optimise a criterion of goodness of fit of the model solution to the data. A mean squared error criterion is frequently used (Berman et al., 1962), but Eykhoff (1963) and Nieman et al. (1971) list a number of alternatives. Metcalf (1973) proposes a median fit such that data lie above or below it with an equal probability.

The algorithms used to iteratively adjust $\bar{k}$ to find the best fit, differ according to the following properties:

1. Their ability to find the same optimum $\bar{k}$, irrespective of any initial guess required;
2. Their ability to indicate the accuracy of the estimated parameters;

3. The computation time required;
4. The convenience of constraining elements of $\bar{k}$;
5. Their ability to estimate from multiple sets of measurements;
6. The convenience with which data may be weighted according to accuracy.

The various methods for the estimation of parameters in nonlinear models are reviewed by Nieman et al. (1971), Aström and Eykhoff (1971), and Bekey (1970). Of note are those methods which take advantage of linearities between the model solution (if it is available analytically) and any of the unknown parameters. Significant computational efficiencies are obtained (Nieman et al., 1971).

3.1.3 Fitting Linear Models

Our general model (equations 3.1 and 3.2, p.37) is linear if $\bar{\mathcal{F}}$ and $\bar{\mathcal{H}}$ are such that each element of $\frac{d\bar{x}(t)}{dt}$ or $\bar{y}(t)$ is a linear sum of elements of $\bar{x}(t)$ and $\bar{r}(t)$. (cf. the example of equations (3.3) and (3.4), p.38). A linear model is described by linear differential equations. Because linear models are special cases of nonlinear models, they are sometimes fitted to data using the methods of section 3.1.2. However, their properties allow application of the special methods extensively reviewed by Bekey (1970), Aström and Eykhoff (1971), and Nieman et al. (1971).

It is well known (Lathi, 1967, p.393) that any output $y(t)$ of a linear model, is related to a single input $r(t)$, by the convolution integral

$$y(t) = \int_0^t i(t-\tau)\, r(\tau)\, d\tau, \qquad (3.17)$$

where i(t) and r(t) are assumed to exist only for $t \geq 0$. In equation (3.17) i(t) is the weighting function or 'impulse response', so named because it is the response of the model to a unit impulse, $\delta(t)$, defined by

$$\delta(t) = 0 \qquad t \neq 0 \tag{3.18}$$

$$\lim_{t \to 0} \delta(t) = \infty \tag{3.19}$$

and

$$\int_{-\infty}^{\infty} \delta(t)\, dt = 1. \tag{3.20}$$

The unit impulse is also called the Dirac delta function (Morse and Feshbach, 1953, p.122).

Equation (3.17) indicates, that for a given r(t), the output of a linear model is completely determined by its impulse response. It is not clear whether the parameters of a linear model should be estimated from the system impulse response, or directly from a measurement of y(t) for an arbitrary input r(t). For some tracer measurements (section 3.3 , p.51) an impulse, or a useful approximation to it, is easily applied to the system in the form of an injection. However, if it is not practical to apply an impulse directly, the impulse response can be calculated from the measured response y(t), to a known input r(t), by solving equation (3.17). The solution of (3.17) for i(t) is called deconvolution, and is discussed in section 3.4.

3.2 COMPARTMENT MODELS

The concept of a compartment is widely used in biological modelling. It is the basis of models of plasma glucose and insulin kinetics (Segre et al., 1973), albumin metabolism (Janes and Carson, 1971), plasma calcium dynamics (Neer et al., 1967), and many other systems.

We use the following definition of a compartment, originally given by Sheppard (1948). A compartment is a quantity of a substance which has uniform and distinguishable kinetics of transformation and transport. It is homogeneous, and may be distinguished by a chemical state, a localization, or both (Rescigno and Segre, 1962). The size of a compartment is the total mass of the substance which it contains.

In CAMET2 (see section 2.2, p.23) the amount of ionized calcium in plasma and some soft tissue, is a compartment. Protein-bound calcium occupies the same physical space as the ionized form (Vaughan, 1970), but it is in a less active chemical state, and is not included in the same compartment.

A compartment is sometimes called a 'pool' (cf. Piessens et al., 1971; Shipley and Clark, 1972, p.1), although Atkins (1969, p.8) states that 'pool' is subject to a wider interpretation. We use only 'compartment'.

A compartment model consists of a number of compartments linked together, so that material may flow from one to another. Material is lost to the environment when it leaves a compartment in the model but does not enter another. For example, CAMET2 (Fig. 2.1, p.23) loses calcium to its environment by urinary excretion from the plasma compartment, and faecal excretion from the final gut compartment. Similarly the stomach compartment gains material from the environment,

in the form of food.

By equating the rate of change of material in a compartment to the difference between the inflow and outflow rates, a model consisting of N compartments is described by

$$\frac{dm_i}{dt} = \sum_{j=1}^{N(i)} r_{ij} + r_{i0} - r_{0i} - \sum_{j=1}^{N(i)} r_{ji} \qquad i = 1,2,\ldots,N \tag{3.21}$$

In (3.21), m_i represents the amount of material in compartment i, r_{ij} is the rate of flow of material from compartment j to compartment i, r_{i0} is the rate at which material enters compartment i from the environment, and r_{0i} is the rate of loss to the environment. $\sum_{j=1}^{N(i)}$ denotes summation with the term for j = i left out. The model (3.21) in general consists of nonlinear differential equations. They become linear if

$$r_{ij} = \mu_{ij} m_j \qquad i \neq j \tag{3.22}$$

and

$$r_{0i} = \mu_{0i} m_i, \tag{3.23}$$

where μ_{ij} and μ_{0i} are constants.

In the next section we find that under certain conditions the use of tracers also results in a linear form of the compartment model (3.21).

3.2.1 Tracers and Compartments

A tracer is a conveniently detected substance whose properties mimic those of another (the tracee) (Atkins, 1969, p.15). Radioactive isotopes are commonly used as tracers. For example calcium-47 is frequently applied in calcium metabolism studies (cf. Neer et al., 1967). Sometimes a dye is used, as in the measurements of circulation transit time reported by

Neufeld (1971).

Following Hearon (1963), q_i is the mass of tracer in compartment i, and a_i its specific activity given by

$$a_i = \frac{q_i}{m_i}. \tag{3.24}$$

Because the tracer mimics the tracee, whose mass is m_i, and is assumed to be uniformly mixed with it, the rate of flow of tracer from a compartment is in proportion to its specific activity, and we find, if $m_i >> q_i$

$$\frac{dq_i}{dt} = \sum_{j=1}^{N(i)} r_{ij} \frac{q_j}{m_j} + S_{i0} - (r_{0i} + \sum_{j=1}^{N(i)} r_{ji}) \frac{q_i}{m_i}$$

$$i = 1,2,\ldots,N. \tag{3.25}$$

In (3.25) S_{i0} is the rate at which tracer flows into compartment i from the environment. The equations (3.25) are linear but with time-varying coefficients $\frac{r_{ij}}{m_j}$, $\frac{r_{ji}}{m_i}$ and $\frac{r_{0i}}{m_i}$. Nieman et al. (1971) refer to techniques for estimating time-varying parameters in a linear model, but we know of no application of these methods to compartment models.

Equation (3.25) is of little practical value, because it is not often possible to measure q_i (cf. Neer et al., 1967); specific activity being the only data available. From (3.24), (3.25) and (3.21) the specific activity in each compartment is given by

$$\frac{da_i}{dt} = \sum_{j=1}^{N(i)} r_{ij} \frac{a_j}{m_i} + \frac{S_{i0}}{m_i} - (r_{0i} + \sum_{j=1}^{N(i)} r_{ij}) \frac{a_i}{m_i}$$

$$i = 1,2,\ldots,N. \tag{3.26}$$

Equation (3.26) is nonlinear unless the tracee is assumed to be in a steady state for which (Atkins, 1969, p.19),

$$\frac{dm_i}{dt} = 0 \qquad i = 1,2,\ldots,N, \tag{3.27}$$

and

$$\frac{dr_{i0}}{dt} = 0 \qquad i = 1,2,\ldots,N. \tag{3.28}$$

The various conditions under which the compartment models of the tracee (3.21), tracer (3.25) and specific activity (3.26) reduce to linear differential equations with constant coefficients, are summarised in Table 3.2, with respect to the model

$$\frac{dy_i}{dt} = \sum_{j=1}^{N} k_{ij}y_j + R_{i0} \qquad i = 1,2,\ldots,N. \tag{3.29}$$

The k_{ij} of equation (3.29) are constant and R_{i0} represents the flow of material into the i^{th} compartment. Table 3.2 also indicates the parameters which can be estimated in the various experimental situations.

We now discuss a special case of the model (3.29):

$$\frac{dy_i}{dt} = \sum_{j=1}^{N} k_{ij}y_j \qquad i = 1,2,\ldots,N, \tag{3.30}$$

where there is no flow of material from the environment. The model (3.30) has zero response unless some of the y_i are given initial values. Commonly (Atkins, 1969, p.20) a measurement is commenced by giving an injection of tracer to the system, at a site corresponding to one of the compartments. For example, the calcium metabolism experiments of Neer et al. (1967) were started with an injection of calcium-47 into the bloodstream. Neer obtained the impulse response directly, but we recall from section 3.1.3 (p.42) that for any arbitrary

Table 3.2

Assumptions under which compartmental models reduce to linear constant coefficient differential equations of the form (3.29)

Model of	Assumption*	y_i	R_{i0}	k_{ij} $(i \neq j)$	k_{ii}
Tracee (3.21) (p.45)	L	m_i	r_{i0}	μ_{ij}	$-(\mu_{0i} + \sum_{j=1}^{N(i)} \mu_{ji})$
Tracer (3.25) (p.46)	S	q_i	S_{i0}	$\frac{r_{ij}}{m_j}$	$-(r_{0i} + \sum_{j=1}^{N(i)} r_{ji})/m_i$
	L	q_i	S_{i0}	μ_{ij}	$-(\mu_{0i} + \sum_{j=1}^{N(i)} \mu_{ji})$
Specific Activity (3.26) (p.46)	S	a_i	$\frac{S_{i0}}{m_i}$	$\frac{r_{ij}}{m_i}$	$-(r_{i0} + \sum_{j=1}^{N(i)} r_{ij})/m_i$
	S and L	a_i	$\frac{S_{i0}}{m_i}$	$\frac{\mu_{ij} m_j}{m_i}$	$-(r_{i0} + \sum_{j=1}^{N(i)} \mu_{ij} m_j)/m_i$

* L means that the tracee has linear dynamics expressed by the constraints (3.22) and (3.23) (p.45).

S means that the tracee is assumed to be at steady state according to (3.27) and (3.28) (p.47).

input R_{i0}, the impulse response can be computed by deconvolution.

3.2.2 Behaviour of Linear Compartment Models

In section 3.1.1 (p.38) we considered a model consisting of two compartments described by

$$\frac{dy_1}{dt} = k_2 y_2 - k_3 y_1 \tag{3.31}$$

$$\frac{dy_2}{dt} = k_1 y_1 - k_4 y_2 \tag{3.32}$$

with

$$y_1(0) = 1 \quad \text{and} \quad y_2(0) = 0. \tag{3.33}$$

The solution for y_1 is

$$y_1 = A_1 e^{-\lambda_1 t} + A_2 e^{-\lambda_2 t} \tag{3.34}$$

where λ_1, λ_2, A_1 and A_2 are given by (3.10), (3.11), (3.12) and (3.13) respectively (p.39). Equation (3.34) describes the dynamics of the material in compartment 1 after an initial injection. The general solution to the model (3.30) is (Coddington and Levinson, 1955, p.78)

$$y_i(t) = \sum_{j=1}^{N} \sum_{\ell=0}^{n-1} A_{ij}\, t^{\ell}\, e^{-\lambda_j t} \qquad i = 1,2,\ldots,N, \tag{3.35}$$

where $-\lambda_j$ is the j^{th} root of the characteristic equation

$$|K - \lambda I| = 0. \tag{3.36}$$

In (3.36) K is the N x N matrix with elements k_{ij}, I is the identity matrix and $|P|$ denotes the determinant of the square matrix P (Elgerd, 1967, p.541). Each root λ_j is repeated n times.

The conditions

$$k_{ij} \geq 0 \qquad i \neq j \tag{3.37}$$

and

$$-k_{ii} \geq \sum_{j=1}^{N(i)} k_{ji} \tag{3.38}$$

are sufficient (but not necessary) to ensure that all the λ_j of (3.35) have positive real parts (cf. Berman and Schoenfeld, 1956). The model is therefore 'stable' so that

$$\lim_{t\to\infty} y_i(t) < \infty \qquad i = 1,2,\ldots,N. \tag{3.39}$$

From Table 3.2 (p.48) we deduce that (3.37) and (3.38) correspond to

$$\begin{aligned} m_i &\geq 0 \\ \mu_{ij} &\geq 0 \\ \mu_{0i} &\geq 0 \\ r_{ij} &\geq 0 \\ \text{and } r_{i0} &\geq 0, \end{aligned} \tag{3.40}$$

which are consistent with the physical interpretation of these quantities.

The conditions under which any λ_j of (3.35) are complex, thereby giving a solution $y_i(t)$ which oscillates, are discussed by Hearon (1963) and Thron (1972). Oscillations can occur only if material leaves a compartment and passes through at least two others, before returning to the first. Thron (1972) and Landahl (1972) report that any oscillations are necessarily of low amplitude and are unlikely to be observed.

3.3 FITTING DECAYING EXPONENTIALS TO DATA

Neer et al. (1967) gave an intravenous injection of calcium-47 to patients, and measured the plasma specific activity at subsequent times. The specific activity was observed to fall to a very low value over a period of 20 days. No oscillations were detected. Neer assumed that the stable calcium was in a steady state during his experiments and that (see Table 3.2, p.48) a linear compartment model of the form (3.30) applies in consequence.

The above is a typical experimental situation using tracers (cf. Atkins, 1969). The solution to the model in the i^{th} compartment is given by (3.35), but because, in a physical system, there is negligible chance that any root λ_j will be repeated, we set n = 1 and write

$$y_i(t) = \sum_{j=1}^{N} A_{ij} e^{-\lambda_j t} \qquad i = 1,2,\ldots,N, \qquad (3.41)$$

where all λ_j are real. Berman and Schoenfeld (1956) show that the parameters k_{ij} of the model (3.30) (p.47) and the A_{ij} and λ_j of (3.41) are related by

$$K A = A \Lambda, \qquad (3.42)$$

where K is an N x N matrix with elements k_{ij}, A is an N x N matrix of elements A_{ij}, and Λ is a diagonal matrix of the λ_j. It is rarely possible to take measurements in all compartments (only the plasma compartment was directly accessible to Neer et al., 1967), and so, because of (3.41) and (3.42), not all the elements of A and K can be estimated. This is the uniqueness problem (cf. section 3.1.1, p.38).

We assume that measurements are available from only one of the N compartments in equation (3.41), and henceforth omit the suffix i, so that

$$y(t) = \sum_{j=1}^{N} A_j e^{-\lambda_j t}. \qquad (3.43)$$

Equation (3.43) is the solution in any compartment of a model for which any of the assumptions of Table 3.2 apply. We now discuss the various ways by which estimates of A_j, λ_j and N respectively are obtained from a measurement of y(t).

Acton (1970, p.252) in an interlude entitled 'What not to Compute', has this to say about attempts to determine the parameters of sums of decaying exponential functions from measured data:

'The answer to this problem lies in the chemical rather than the computer laboratory, and the sooner the hopeful innocent can be sent there and away from the computer room the better off everyone will be.'

The difficulty of fitting exponential functions implied by Acton's comment, is mainly due to their nonorthogonality over any time interval (Lanczos, 1956, p.272).

3.3.1 Review of Existing Methods

Consider the example

$$y(t) = e^{-t} + e^{-0.1t}. \qquad (3.44)$$

We note that as t becomes large y(t) depends almost entirely on the second (slower decaying) term $e^{-0.1t}$. Further, if we plot the logarithm of y(t) (ln y(t)), as a function of time, we expect the graph to be a straight line of slope -0.1 for

large t, as shown in Fig. 3.3.

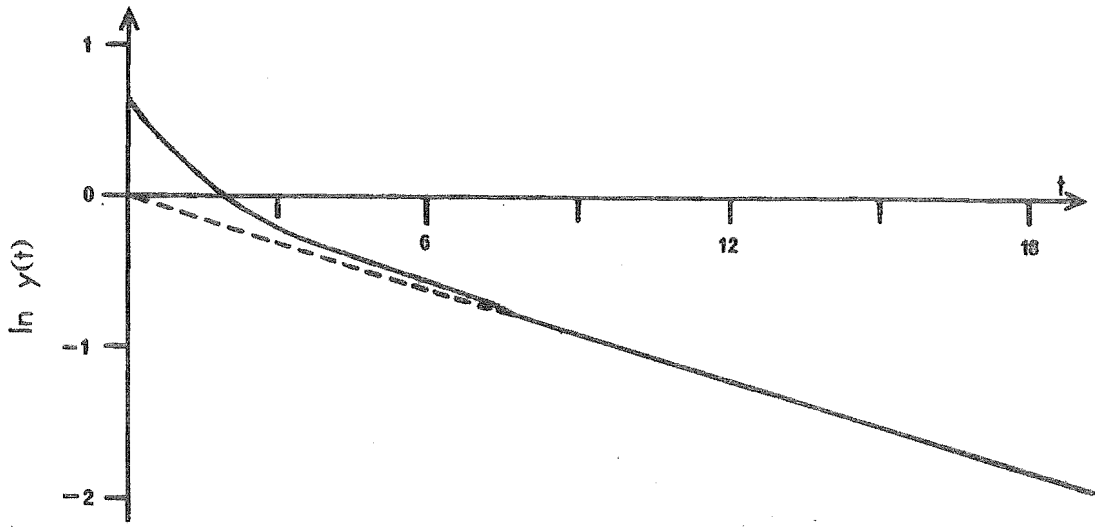


Fig. 3.3 Principle of the peeling method for the example of equation (3.44)

This forms the basis of the classical peeling method for estimating the parameters of decaying exponentials. When the slowest decaying component has been estimated as in Fig. 3.3, the straight line corresponding to it is extrapolated back to $t = 0$ and subtracted out. The process is repeated for the next slowest decaying term. Mancini and Pilo (1970) describe a computer implementation of the peeling method, in which a statistical criterion is used to isolate the final linear portion of the logarithmic graphs. The number of exponentials N is the total number of terms identified before the data is exhausted. Perl (1960) proposes an alternative peeling approach in which $\frac{dy(t)}{dt}$ is plotted against $y(t)$. The slope of the final linear section of the graph again gives the slowest decaying exponent. The corresponding term is then subtracted from the data, and the process repeated for the next

slowest term. The values of $\frac{dy(t)}{dt}$ are obtained by differentiating the measured data.

Parameters of decaying exponentials are frequently estimated using the nonlinear optimisation methods referred to in section 3.1.2 (cf. Piessens et al., 1971; Berman et al., 1962). Various nonlinear optimisation algorithms are discussed by Brown (1972), Bekey (1970), Wilson (1970) and Davis and Ottaway (1972). Lemaitre and Malengé (1971) use the linear relationship between the A_j and y(t) of (3.43) (p.52) to achieve significant computational savings for a least squares criterion of goodness of fit (see section 3.1.2, p.42). They calculate the number of exponentials by repeating the estimation with increasing N, until a statistical test indicates that the fit is not significantly improved. To estimate parameter accuracy it is frequently assumed that all parameters being estimated are linearly related to y(t) in the region of best fit (cf. Berman et al. 1962). Glass and de Garreta (1967) estimate parameter accuracy by a Monte Carlo method. Having obtained the best fit, they add random error, approximating the experimental error, to the fitted curve. The result is used as data to repeat the parameter estimation. They carry out the process with various random error sequences, until the statistics of each parameter have been estimated.

Prony's method, first published in 1795, is the basis of a number of procedures for fitting exponential functions (cf. Cornell, 1962; Rescigno and Segre, 1962; Parsons, 1968). Tuttle (1971) describes the historical background to this method, and some of its applications. Equation (3.43) satisfies the difference formula

$$c_0 y(t) + c_1 y(t+d) + c_2 y(t+2d) + \ldots + y(t+Nd) = 0. \tag{3.45}$$

Letting

$$\xi_j = e^{-\lambda_j d}, \tag{3.46}$$

substitution of (3.43) into (3.45) yields

$$c_0 + c_1 \xi + c_2 \xi^2 + \ldots + \xi^N = 0, \tag{3.47}$$

where the N roots of (3.47) are $\xi = \xi_1, \xi_2, \ldots, \xi_N$. The constants $c_0, c_1, \ldots, c_{N-1}$ of (3.47) are obtained by solving the linear equations

$$\begin{aligned} y_1 c_0 + y_2 c_1 + \ldots + y_N c_{N-1} &= -y_{N+1} \\ y_2 c_0 + y_3 c_1 + \ldots + y_{N+1} c_{N-1} &= -y_{N+2} \\ &\vdots \\ y_N c_0 + y_{N+1} c_1 + \ldots + y_{2N-1} c_{N-1} &= -y_{2N}, \end{aligned} \tag{3.48}$$

where

$$y_i = y(t + (i-1)d). \tag{3.49}$$

When the λ_j have been found by (3.46) from the roots of the polynomial (3.47), the A_j are determined by a further matrix inversion. The order of the model, N, is the rank of the matrix of y values in (3.48). Myhill et al. (1965) estimate N by increasing the number of equations in (3.48), until the matrix inversion becomes inaccurate. Parsons (1968) calculates the determinant of the equations (3.48) and statistically tests its change in value, as N is increased.

Prony's method as described above is very sensitive to small errors in the data (cf. Lanczos, 1956, p.275). To reduce the effects of data error, Myhill et al. (1965) formulate

(3.48) as an overdetermined set of equations which are solved by linear least squares. However, theirs is not a standard least squares problem because the values in error (i.e. measurements of y(t)) occur on both sides of the equation. They weight the data in an attempt to compensate.

Diamessis (1972) proposes to estimate the parameters of exponential functions by expressing (3.43) (p.52) as an N^{th} order differential equation with constant coefficients and appropriate initial conditions. The coefficients are calculated using a linear least squares procedure, set up after repeated integration of the data. The λ_j are the roots of an N^{th} order polynomial as in Prony's method.

The transform

$$\Upsilon(s) = \int_0^\infty st\ y(t)\ \sin st\ dt \tag{3.50}$$

is tested by Brownell and Callahan (1963). Substituting the exponential sum (3.43) into (3.50) yields

$$\Upsilon(s) = \sum_{j=1}^{N} \frac{2A_j\lambda_j s^2}{(s^2+\lambda_j^2)^2} \tag{3.51}$$

which contains peaks at $s = \lambda_j$, provided the λ_j are sufficiently distinct. Brownell and Callahan (1963) compute $\Upsilon(s)$ from the data y(t). The λ_j are the values of s at which peaks appear in $\Upsilon(s)$, and N is the number of peaks.

An approach requiring the inversion of the Laplace transform is described by Pizer et al. (1969). They express y(t) as a Laplace transform

$$y(t) = \int_0^\infty Y(\alpha)\ e^{-\alpha t}\ d\alpha. \tag{3.52}$$

Inspection of (3.52) and the exponential sum (3.43) yields

$$Y(\alpha) = \sum_{j=1}^{N} A_j \, \delta(\alpha-\lambda_j), \qquad (3.53)$$

where $\delta(\alpha-\lambda_j)$ is the Dirac delta function defined in the glossary. In chapter 4 we discuss methods of solution of (3.52) for $Y(\alpha)$ given a measurement of $y(t)$. The exponents λ_j are given by the position of the peaks in $Y(\alpha)$, and N by the number of peaks. In principle the method is capable of separating exponential functions with nearly equal exponents, and, by the 'peakiness' of $Y(\alpha)$, capable of testing the hypothesis that the data actually does consist of a sum of decaying exponentials.

3.3.2 Conclusions

With the exception of the method described by Diamessis (1972), the various techniques for fitting decaying exponential functions described in section 3.3.1 have been subject to computational tests. We draw conclusions from these tests as reported by various authors.

Typically, ideal data is generated from (3.43) with chosen values of A_j, λ_j and N. Controlled random error is superimposed on the data, which is then analysed by the particular technique being tested. It is of interest to know how well the estimates A_j, λ_j and N conform to the original A_j, λ_j and N respectively.

Where the number of compartments is not known a priori (as in, for example, the calcium tracer tests of Neer et al., 1967), it would be desirable to correctly estimate N from the data. We conclude that this cannot be done. Table 3.4 lists the approximate conditions under which the various methods

Table 3.4

Conditions for Incorrect Estimation of N

Method	Standard deviation of error (% of value of each data point)	Minimum exponent ratio R	Reference
Peeling	1%	1.43	Mancini and Pilo (1970)
Nonlinear Optimisation	2%	3	Piessens et al. (1971)
Prony[1]	2%	2	Myhill et al. (1965)
Transform (3.50) (p.56)	0	4	Brownell and Callahan (1963)
Inverse Laplace Transform	0	1.5	Wilson (1970)[2]

[1] This is a least squares version of Prony's method in which preliminary smoothing of the data is also carried out.

[2] This result is obtained from data available over a very large time interval, but it is only an indication, because the resolving power of the method depends in a complex way on the exponents λ_j, the coefficients A_j, the data error and, the time interval over which the data is available (cf. Wilson, 1970; Pizer et al., 1969; and section 4.2.1).

have estimated N incorrectly. In Table 3.4, R is defined as the ratio of nearest exponents such that R > 1, and we expect the estimate of N to deteriorate further from that given in the table as R approaches unity (i.e. as the exponents become more indistinct). The data error given in Table 3.4 is normally distributed, with standard deviation a fixed percentage of the value of each datum. More accurate experimental data is unlikely to be available (Piessens et al., 1971).

Unless N is already known, it is usual to find the smallest value which allows satisfactory representation of the data (cf. Neer et al., 1967). This gives the 'minimal model' referred to by Kalman (1968).

Even if the correct N is known, the authors cited in Table 3.4 report instances of A_j and λ_j estimates being in error by 100% when data error is as low as 5%. No method gives estimates significantly more correct than any other. We note that in all cases, the estimates which are obtained, give a fit to the data which looks entirely satisfactory when plotted on a graph.

We therefore choose on the basis of convenience of use, and conclude from Table 3.5 that the nonlinear optimisation methods are to be preferred for their flexibility.

Glass and de Garreta (1967) show by experiment, that the assumption of linearity in the region of best fit can lead to the overestimation of the errors in A_j and λ_j by up to 100%. Their alternative Monte Carlo procedure is to be preferred as it is more soundly based, but the additional computing required may not be justified.

Table 3.5

Properties of Exponential Function Identification Methods

Method	Data Weighting	Data Constraints	Useful for N fixed?	Parameter Error Estimates	Initial Guess Required?	Nature of Computation
Peeling	specified by user	none	no	not available	yes[3]	iterative
Nonlinear Optimisation	specified by user	none	yes[1]	usually available	yes	iterative
Prony[2]	cannot be specified	required at equal time intervals	yes[1]	not available	no	direct
Inverse Laplace Transform	cannot be specified	required at equal intervals of log time	no	not available	no	direct
Transform (3.50) (p.56)	cannot be specified	required at equal time intervals	no	not available	no	direct

[1] If the data does not conform to the uniqueness and consistency conditions discussed in section 3.1.1, computational difficulties may occur for a particular N value.

[2] The implementation of Prony's method by Myhill et al. (1965) incorporates a preliminary data smoothing procedure.

[3] This is an unnecessary restriction in the implementation by Mancini and Pilo (1970).

In summary, we conclude that the number of exponentials (compartments) cannot be determined from the data. We prefer a nonlinear optimisation procedure for the estimation of decaying exponentials but we must expect large uncertainties in the estimates A_j and λ_j.

3.4 DECONVOLUTION OF DATA

In section 3.1.3 (p.42) we referred to the relationship

$$y(t) = \int_0^t i(t-\tau)\, r(\tau)\, d\tau \qquad (3.55)$$

between the input $r(t)$ and output $y(t)$ of a linear model. The function $i(t)$ is the impulse response and is determined from (3.55) by deconvolution, if $y(t)$ and $r(t)$ are known.

Equation (3.55) is a special case of the convolution integral given by Lathi (1967, p.148) as

$$f(x) = \int_{-\infty}^{\infty} g(x-\xi)\, h(\xi)\, d\xi. \qquad (3.56)$$

If $g(x)$ and $h(x)$ are causal, that is

$$\left.\begin{aligned} g(x) &= 0 \\ h(x) &= 0 \end{aligned}\right\} \quad \text{for } x < 0, \qquad (3.57)$$

equation (3.56) reduces to the form (3.55). In the many deconvolution problems which arise in science and engineering, x in (3.56) represents either a spatial or a time variation. The determination of an impulse response is an example of the latter, and in section 7.4 we describe a further example, where the input is calculated from output and impulse response measurements. Optical image processing (which may involve deconvolution in two dimensions) is a situation in

which x represents a spatial variation (IEEE, 1972), as is the correction for instrument broadening effects in X-ray diffraction (section 6.3).

In what follows we distinguish two types of problem.

Type A: Functions g(x) and f(x) are given for all x at which they have values significantly different from zero. The data g(x) and f(x) are therefore 'complete'. Type A problems occur for example in image processing (IEEE, 1972) and X-ray diffraction (section 6.3).

Type B: Either g(x) or h(x) is of large (perhaps infinite) extent so that f(x) cannot be measured throughout its entire interval of existence. Equation (3.55) often implies a Type B problem, because the impulse response usually exists over a longer time interval than that for which measurements of y(t) are available. The calcium absorption test using two radioactive isotopes is a Type B problem (section 7.4).

3.4.1 Review of Deconvolution Methods

Jones and Misell (1970) point out that deconvolution is an unstable process, and to counter the effects of data error, some sort of approximation is required. Methods vary according to the nature of the approximation.

In the simplest deconvolution procedures, the integral (3.56) is approximated by a sum, and the resulting linear equations are solved. Various approximation formulae are used (cf. Branston and Read, 1972; Louër and Weigel, 1969). These methods are mainly applied to Type B problems (Bekey, 1970), because they enable h(x) to be calculated throughout the entire interval in which f(x) is known. Preliminary smoothing of the data is carried out (Branston and Read, 1972) or sometimes the data are fitted by a parametric equation

(Szymendera et al., 1972). The convolution integral (3.56) is approximated by a sum in Type A problems, in which an initial guess for h(x) is iteratively adjusted until f(x) is best represented (Jansson, 1970). Only a few iterations are possible before the result sometimes diverges (Jones and Misell, 1970), but this method does allow constraints to be imposed on the smoothness and positivity of h(x).

The discrete Laplace transform, or Z transform, has been applied to Type B problems, as it is mathematically equivalent to approximating the convolution integral (3.56) by a summation (cf. Neufeld, 1971; Branston and Read, 1972; Hli, 1971). When the causal conditions (3.57) apply to (3.56) (p.61), the approximation of the convolution integral by a sum is equivalent to considering f(x) to exist only at a set of sample points distance d apart. That is

$$f(x) = \sum_{n=0}^{\infty} f_n \delta(x-nd). \tag{3.58}$$

Defining the Laplace transform of f(x) as (cf. Morse and Feshbach, 1953, p.467)

$$F_\ell(s) = \int_0^\infty f(x)\, e^{-sx}\, dx , \tag{3.59}$$

we get

$$F_\ell(s) = \sum_{n=0}^{\infty} f_n\, e^{-snd}. \tag{3.60}$$

We let

$$z = e^{sd} \tag{3.61}$$

and obtain the Z transform of f(x) given by

$$Z\{f(x)\} = \sum_{n=0}^{\infty} f_n z^{-n} \tag{3.62}$$

We note that $Z\{f(x)\}$ is a polynomial in z, as are $Z\{g(x)\}$, and $Z\{h(x)\}$, defined similarly. If h(x) and g(x) are causal (3.57), taking the Laplace transform of g(x) and h(x) in (3.56) yields

$$F_\ell(s) = G_\ell(s)\, H_\ell(s) \tag{3.63}$$

or

$$Z\{f(x)\} = Z\{g(x)\}\, Z\{h(x)\}. \tag{3.64}$$

Therefore $Z\{h(x)\}$ is obtained by the long division of the polynomial $Z\{f(x)\}$ by the polynomial $Z\{g(x)\}$. Neufeld (1971) improved the stability of the method by smoothing the coefficients of the remainder polynomial at each step in the division.

Type A problems are mostly solved using methods based on the Fourier transform defined by (cf. Morse and Feshbach, 1953, p.453)

$$F(u) = \int_{-\infty}^{\infty} f(x)\, e^{j2\pi ux}\, dx \tag{3.65}$$

and its inverse

$$f(x) = \int_{-\infty}^{\infty} F(u)\, e^{-j2\pi ux}\, du. \tag{3.66}$$

The transforms G(u) of g(x), and H(u) of h(x) are defined similarly. Taking the Fourier transform of f(x) in (3.56)

(p.61) yields

$$F(u) = G(u)\ H(u). \tag{3.67}$$

Therefore deconvolution is carried out by calculating F(u) and G(u) from f(x) and g(x) using (3.65), computing

$$H(u) = \frac{F(u)}{G(u)}, \tag{3.68}$$

and applying the inverse transform (3.66) to determine h(x). There are two well-known difficulties with this inverse filtering procedure. Firstly, G(u) may be zero for some u, when, due to data error, F(u) is not, thus causing H(u) to be undefined. Secondly, data error frequently causes G(u) to decrease faster than F(u) as u becomes large, so that the H(u) in the inverse integral (3.66) does not converge. If the computations are done on a digital computer as in section 8.1 (they can also be done optically, cf. Goodman, 1968), the sampling of F(u) and G(u) can be arranged to avoid points at which G(u) is zero and so solve the first problem, despite the theoretical objections of Sondhi (1972). The second difficulty is overcome by applying a filter function to H(u) which ensures that it becomes well behaved for large u (Helstrom, 1967). In some methods f(x), g(x) and h(x) are represented by a linear sum

$$f(x) = \sum_{m=1}^{M} F_m\ \phi_m(x), \tag{3.69}$$

where the $\phi_m(x)$ are functions chosen so that only a few terms in the expansion (3.69) are required to represent the data. Hossfeld (1968) and Berry (1947) use Hermite functions for

$\phi_m(x)$, whereas Moore (1968) uses Fourier series.

Fourier methods are not popular for Type B problems because F(u) must be in error as it is calculated from a truncated f(x). Silverman and Pearson (1973) and Neufeld (1971) apply the Fourier method iteratively to extrapolate the truncated f(x) so that correct deconvolution is obtained.

The methods cited above and their variations have been tailored by various authors to suit particular deconvolution problems. However, in every case, the quality of the original data is assessed mainly on the basis of the result of the deconvolution. In chapters 5, 6 and 7 we develop and use consistency conditions on the data f(x) and g(x).

CHAPTER 4

FITTING EXPONENTIAL FUNCTIONS TO DATA WITH THE INVERSE LAPLACE TRANSFORM

In section 3.3.1 (p.56) we mentioned a method for estimating the parameters of exponential functions, in which the data $y(t)$ is expressed as the Laplace transform of a function $Y(\alpha)$ by

$$y(t) = \int_0^\infty Y(\alpha)\ e^{-\alpha t}\ d\alpha. \qquad (4.1)$$

If $y(t)$ is the exponential sum

$$y(t) = \sum_{i=1}^{N} A_i\ e^{-\lambda_i t}, \qquad (4.2)$$

inspection of (4.1) yields

$$Y(\alpha) = \sum_{i=1}^{N} A_i \delta(\alpha-\lambda_i). \qquad (4.3)$$

This formulation has, in principle, the following useful characteristics:

1. The number of exponentials N is simply observed as the number of peaks in $Y(\alpha)$.
2. Exponentials with nearly equal exponents can be resolved.
3. The form of $Y(\alpha)$ gives a check on the consistency of the hypothesis that the data does actually represent an exponential sum.

When $y(t)$ describes measured data, the best $Y(\alpha)$ that we can obtain is never such that the above three conditions can be fully realised.

The solution of (4.1) for $Y(\alpha)$ is the inversion of the Laplace transform, and has an extensive literature in electrical network analysis (cf. Piessens, 1971), and time-dependent transport calculations (cf. Renken and Biggs, 1972). The methods for numerical inversion of the Laplace transform described in the following sections require the introduction of the complex variable

$$p = t + j\tau \tag{4.4}$$

so that (4.1) may be rewritten

$$y(p) = \int_0^\infty Y(\alpha)\ e^{-\alpha p}\ d\alpha. \tag{4.5}$$

4.1 NUMERICAL INVERSION OF THE LAPLACE TRANSFORM

The formula for the inversion of the Laplace transform (4.5) is given by Morse and Feshbach (1953, p.468) as

$$Y(\alpha) = \frac{1}{j2\pi} \int_{Br} y(p)\ e^{\alpha p} dp \qquad Re\{\alpha\} > 0. \tag{4.6}$$

In (4.6)

$$j = \sqrt{-1} \tag{4.7}$$

and Br represents the Bromwich contour, which is a line of infinite extent, parallel to the imaginary ($j\tau$) axis, but lying to the right of all singularities of $y(p)$. We note that to calculate $Y(\alpha)$ using equation (4.6) it is necessary to know $y(p)$ for complex values of p along Br. Because our data are a function of only the real variable t, we must find a representation of it in the complex plane.

If an analytical expression is available for $y(p)$, such as arises in electrical network problems (cf. Silverberg, 1970),

$Y(\alpha)$ is computed directly from (4.6). The complexity of $y(p)$ sometimes prevents an analytical solution, and in such cases a number of techniques (known as quadrature formulae) are used to approximate the integration numerically (cf. Zakian, 1970; Piessens, 1971).

Bellman et al. (1966, p.17) treat (4.5) as an integral equation, and show that its solution for $Y(\alpha)$ is unstable with respect to small perturbations in $y(t)$. They use a different quadrature approximation than Cook et al. (1967), but in each case the matrix to be inverted is 'poorly conditioned' (cf. Westlake, 1968, p.88), and a solution $Y(\alpha)$ is obtained only by applying smoothness constraints. Renken and Biggs (1972) represent $Y(\alpha)$ in terms of special functions, and use a non-linear optimisation procedure to determine the unknown parameters in this representation. For our problem, the latter procedure is equivalent to exponential fitting using the nonlinear optimisation methods mentioned in section 3.3.1 (p.54). We expect $Y(\alpha)$ to contain sharp peaks, and so, not wishing to apply a smoothness constraint, we conclude that the integral equation approach is unsuitable for our purposes.

Our data $y(t)$ are likely to be a set of samples taken at various times throughout the finite interval $0 \leqslant t \leqslant T_y$ (cf. Neer et al., 1967). It can only be described as analytic in the sense that the measured values are assumed to be samples of an analytic function (e.g. a sum of exponentials). We make this assumption. Analytic continuation is the process by which a function which has a valid representation in a given region of the complex plane, is given a representation in an adjacent region (Morse and Feshbach, 1953, p.389). The function must be analytic in both regions. In our case $y(t)$

is given on the real line, and we wish to represent it along Br.

Notionally, y(t) exists for

$$0 \leqslant t < \infty, \tag{4.9}$$

and so we might consider the analytic continuation of y(t) from the interval $0 \leqslant t \leqslant T_y$ in which it is actually measured, to the rest of the real line $T_y \leqslant t < \infty$. The inevitable measurement error in y(t) means that it has only a finite number of physically meaningful derivatives. This precludes any successful extrapolation using a classical Taylor series approach (Morse and Feshbach, 1953, p.392). An alternative is to obtain a functional representation for y(t) in the interval $0 \leqslant t \leqslant T_y$, and use it for analytic continuation to the interval $T_y \leqslant t < \infty$. Our main objective is to derive a functional representation (i.e. a sum of exponentials) for y(t) in the interval $0 \leqslant t \leqslant T_y$, and so we do not attempt the extrapolation of measured data until after the parameters of the exponential functions have been estimated.

4.2 ANALYTIC CONTINUATION USING INTEGRAL TRANSFORMS

We define

$$f(\xi) = \int_C y(t)\, K(\xi,t)\, dt, \tag{4.10}$$

which is an integral transform formula. We postulate that (4.10) has the inverse

$$y(p) = \int_{C'} f(\xi)\, K^{-1}(p,\xi)\, d\xi \tag{4.11}$$

where C and C' are contours in the complex p-plane and the complex ξ-plane respectively. The first transformation (4.10), with kernel function $K(\xi,t)$, is an operation on the measurement $y(t)$, and so we require that the contour C be that portion of the real line t, over which $y(t)$ is known (i.e. $0 \leq t \leq T_y$). However, in all known transforms, C is infinitely long, and we shall observe how knowing $y(t)$ only for $0 \leq t \leq T_y$ reduces the accuracy of the procedure. For the reverse transformation (4.11), with kernel function $K^{-1}(p,\xi)$ (note that K^{-1} denotes a new function, and is not $1/K$), the contour C' may be the real line as for Fourier, Hankel and Hilbert transforms, or a Bromwich contour as for Laplace and Mellin transforms (cf. Morse and Feshbach, 1953, p.942).

To be useful for our analytic continuation, any transform with kernels $K(\xi,t)$ and $K^{-1}(p,\xi)$ and contours C and C' must be such that the integrals (4.10) and (4.11) exist. Specifically we require, at least that

$$\lim_{t\to\infty} y(t)\ K(\xi,t) = 0 \qquad (4.12)$$

for all ξ on C', and

$$f(\xi)\ K^{-1}(p,\xi) \to \text{a constant} \qquad (4.13)$$

as ξ approaches the extremities of the contour C', for all p on the Bromwich contour of the Laplace inversion formula (4.6).

Of the integral transforms mentioned above, only the Mellin transform satisfies the conditions (4.12) and (4.13) when $y(t)$ is assumed to decay exponentially. For the Mellin transform, C is the contour $0 \leq t < \infty$, C' is a Bromwich contour, and (cf. Morse and Feshbach, 1953, p.469)

$$K(\xi,t) = t^{\xi-1}, \tag{4.14}$$

and

$$K^{-1}(p,\xi) = p^{-\xi}/j2\pi. \tag{4.15}$$

The formulae (4.10) and (4.11) become

$$f(\xi) = \int_0^\infty y(t)\ t^{\xi-1}\ dt \tag{4.16}$$

and

$$y(p) = \frac{1}{j2\pi} \int_{Br} f(\xi)\ p^{-\xi}\ d\xi. \tag{4.17}$$

We define ξ to be the complex variable

$$\xi = \eta + j\zeta, \tag{4.18}$$

where η and ζ are real.

To demonstrate analytic continuation with the Mellin transform let

$$y(t) = e^{-\lambda_1 t} \tag{4.19}$$

and so

$$f(\xi) = \int_0^\infty e^{-\lambda_1 t}\ t^{\xi-1}\ dt. \tag{4.20}$$

If we set $\eta \geqslant 1$, so that the integrand of (4.20) behaves well at $t = 0$, we find

$$f(\xi) = \frac{\Gamma(\xi)}{\lambda_1^{\xi}}, \tag{4.21}$$

where $\Gamma(\xi)$ is the gamma function (cf. Abramowitz and Stegun 1965, p.255). The function $\Gamma(\xi)$ has a pole at the origin, but is otherwise analytic in the right half plane. Therefore, the restriction $\eta \geqslant 1$ is sufficient to ensure that the Bromwich contour of (4.17) is to the right of any poles of

$f(\xi)$. We choose $\eta = 1$, and for (4.17) obtain

$$y(p) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Gamma(1 + j\zeta)(\lambda_1 p)^{-(1+j\zeta)} \, d\zeta. \tag{4.22}$$

We choose the imaginary ($j\tau$) axis as the Bromwich contour for the Laplace inversion integral (4.6), so that (4.22) reduces to

$$y(j\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Gamma(1+j\zeta)(j\tau\lambda_1)^{-(1+j\zeta)} d\zeta, \tag{4.23}$$

the integrand of which satisfies (4.13) because (cf. Abramowitz and Stegun, 1965, p.256)

$$|\Gamma(1+j\zeta)|^2 = \frac{\pi\zeta}{\sinh \pi\zeta}. \tag{4.24}$$

Thus, as expected,

$$y(j\tau) = e^{-j\lambda_1\tau}, \tag{4.25}$$

is the analytic continuation of $e^{-\lambda_1 t}$ on to the imaginary axis, and we obtain

$$Y(\alpha) = \delta(\alpha-\lambda_1) \tag{4.26}$$

using the inversion formula (4.6).

To estimate the parameters of exponential functions, the above procedure requires the calculation of $f(\xi)$ using (4.16), the calculation of $y(p)$ from (4.17) and the calculation of $Y(\alpha)$ from (4.6). However, the number of transformations required can be reduced to two by the following procedure. We take the Mellin transform

$$f(\xi) = \int_0^{\infty} y(t) \, t^{\xi-1} \, dt \tag{4.27}$$

and substitute equation (4.1) (p.67) for y(t) yielding

$$f(\xi) = \int_0^\infty \int_0^\infty Y(\alpha)\ e^{-\alpha t}\ d\alpha\ t^{\xi-1}\ dt. \tag{4.28}$$

Because the limits of integration are independent, we may change the order of the integrals in (4.28) giving

$$f(\xi) = \int_0^\infty Y(\alpha) \int_0^\infty e^{-\alpha t}\ t^{\xi-1}\ dt\ d\alpha, \tag{4.29}$$

which reduces to

$$f(\xi) = \int_0^\infty Y(\alpha)\ \frac{\Gamma(\xi)}{\alpha^{\xi}}\ d\alpha. \tag{4.30}$$

Replacing ξ with $1-\xi$ in (4.30) we recognise that

$$\frac{f(1-\xi)}{\Gamma(1-\xi)} = \int_0^\infty Y(\alpha)\ \alpha^{\xi-1}\ d\alpha \tag{4.31}$$

is in the form of a Mellin transform, and so $Y(\alpha)$ is obtained using the inversion formula:

$$Y(\alpha) = \frac{1}{j2\pi} \int_{Br} \frac{f(1-\xi)}{\Gamma(1-\xi)}\ \alpha^{-\xi} d\xi, \tag{4.32}$$

where we may take Br to be the imaginary ($j\zeta$) axis. Therefore, we first calculate

$$f(1+j\zeta) = \int_0^\infty y(t)\ t^{j\zeta}\ dt, \tag{4.33}$$

and then compute

$$Y(\alpha) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{f(1+j\zeta)}{\Gamma(1+j\zeta)}\ \alpha^{j\zeta}\ d\zeta. \tag{4.34}$$

The formulae (4.33) and (4.34) are the same as derived by Pizer et al. (1969) except they let

$$\alpha = e^{-\beta} \tag{4.35}$$

and

$$t = e^{\gamma} \tag{4.36}$$

which convert (4.33) and (4.34) to their equivalent Fourier transforms

$$f(1+j\zeta) = \int_{-\infty}^{\infty} e^{\gamma} y(e^{\gamma})\ e^{j\zeta\gamma} d\gamma \tag{4.37}$$

and

$$Y(e^{-\beta}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{f(1+j\zeta)}{\Gamma(1+j\zeta)}\ e^{-j\beta\zeta} d\zeta. \tag{4.38}$$

4.2.1 Performance of the Mellin Transform Method

To estimate the parameters of exponential functions with the Mellin transform, we use the Fourier transform version of the method given by equations (4.37) and (4.38). The computations are made considerably more efficient than those reported by Pizer et al. (1969) and Gardner et al. (1959), by using the fast Fourier transform (FFT) algorithm (section 8.1). We note from equation (4.37), that the data y(t) must be available at uniform intervals of γ (i.e. ln t), for the FFT to be applied.

Figures 4.1 and 4.2 show the functions

$$R(\zeta) = f(1+j\zeta)/\Gamma(1+j\zeta), \tag{4.39}$$

and $Y(e^{-\beta})$, computed from the ideal data

$$y(t) = e^{-t} \qquad 0 \leqslant t \leqslant 15. \tag{4.40}$$

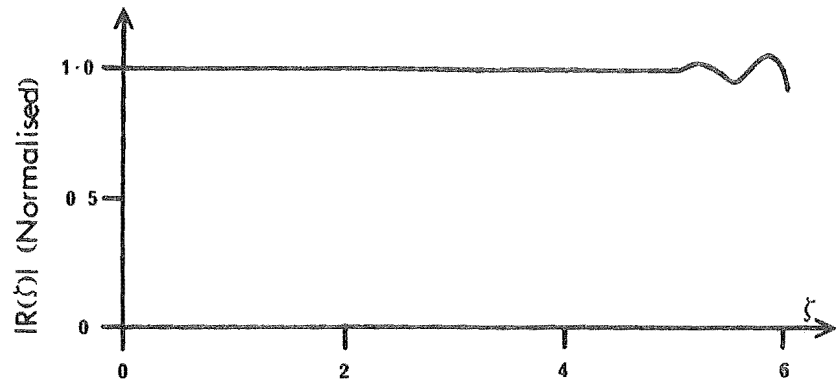


Fig. 4.1 $|R(\zeta)|$ computed from the ideal data of equation (4.40)

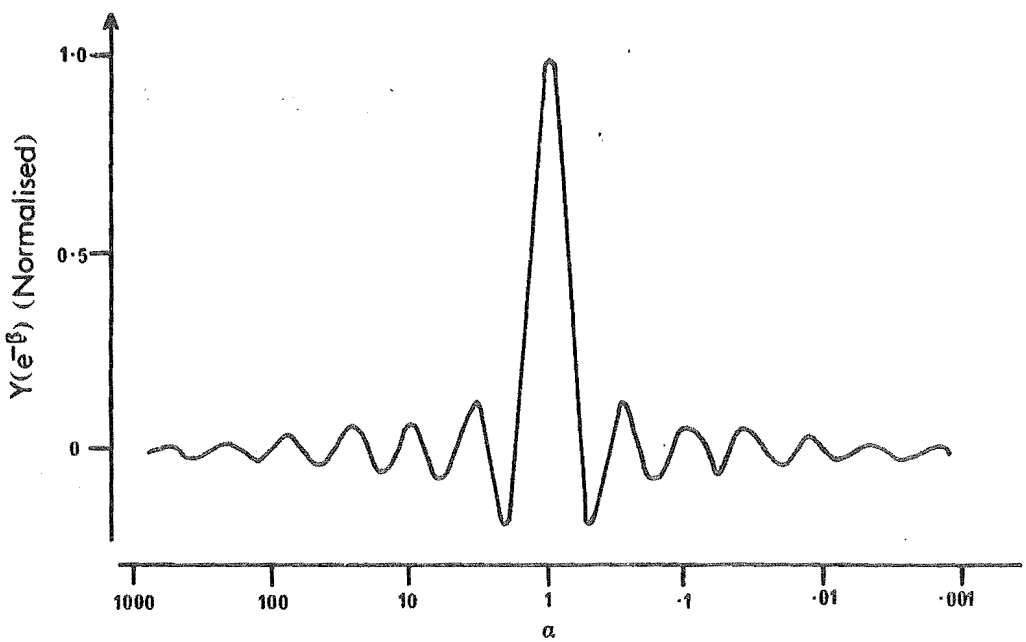


Fig. 4.2 $Y(e^{-\beta})$ computed from the ideal data of equation (4.40)

Substituting for y(t) in equation (4.37) shows that $R(\zeta)$ is theoretically constant for all ζ. Therefore, the integrand in (4.38) does not converge as $|\zeta| \to \infty$, and so it is necessary to truncate the integral (4.38) at some

$$|\zeta| = \zeta_{max}. \qquad (4.41)$$

We note in Fig. 4.1 that at its extremities $R(\zeta)$ deviates from a constant value, despite the ideal nature of the data. Replacing the limits of the integral (4.38) with $-\zeta_{max}$ and ζ_{max}, we find for the example (4.40) that

$$Y(e^{-\beta}) = \frac{2 \sin \beta \zeta_{max}}{\beta} , \qquad (4.42)$$

which is the form of the computed result Fig. 4.2. The vertical scale in Figures 4.1 and 4.2 is normalised. In this and other examples we are concerned primarily with finding the exponents, which are given by the positions of the peaks of $Y(e^{-\beta})$, and so we plot only a normalised amplitude.

Truncation of $R(\zeta)$, before calculating the inverse transform (4.38), increases the width of the peaks in $Y(e^{-\beta})$, and therefore reduces the resolution of the method. The amplitude of the auxilliary peaks or 'sidelobes' associated with the function $\frac{\sin \beta\zeta_{max}}{\beta}$ is reduced by smoothly tailing off $R(\zeta)$ to zero for $|\zeta| \geq \zeta_{max}$. Fig. 4.3 shows the 'cosine bell' function of width 3.7 (section 8.1) applied to the $R(\zeta)$ of the example (4.40).

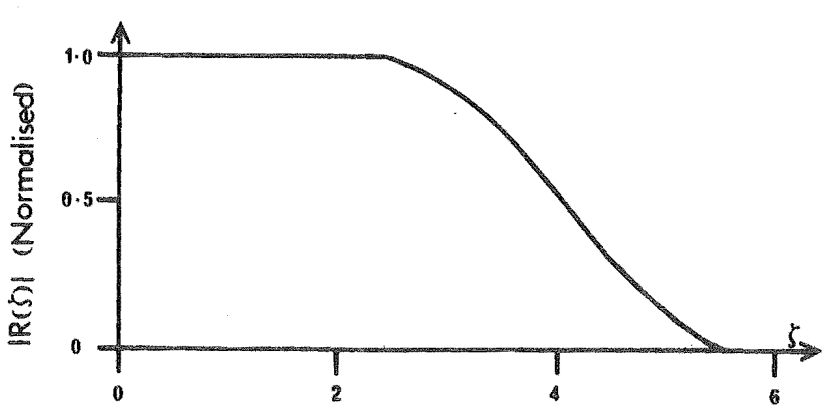


Fig. 4.3 Cosine bell weighting function applied to the $R(\zeta)$ of Fig. 4.1.

The consequent reduction in sidelobes is apparent from Fig. 4.4 (cf. Fig. 4.2, p.76).

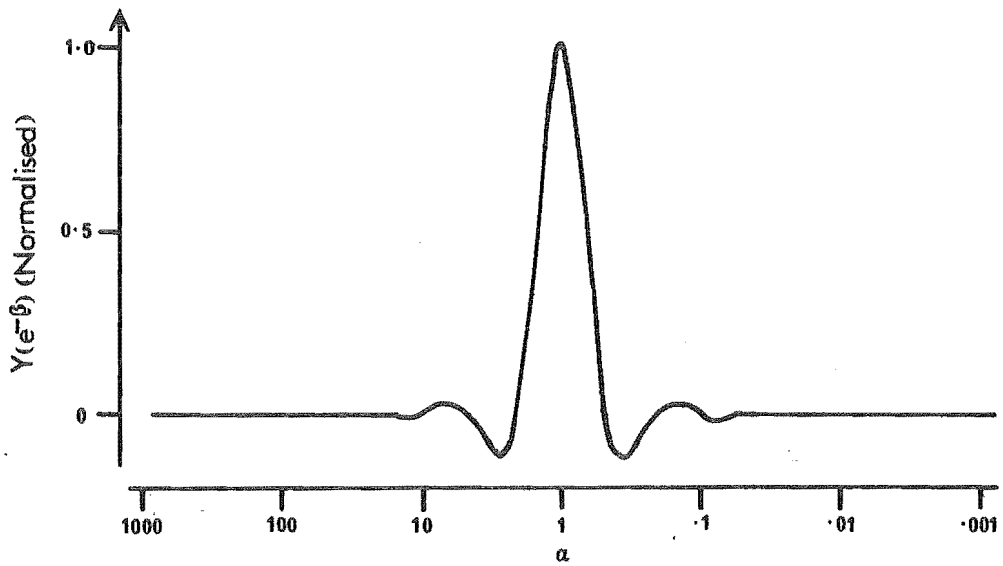


Fig. 4.4 $Y(e^{-\beta})$ computed from the ideal data of equation (4.40) but with $R(\zeta)$ weighted as in Fig. 4.3.

Substitution of the exponential sum (4.2) into equation (4.37) yields

$$Y(e^{-\beta}) = 2 \sum_{i=1}^{N} \frac{A_i \sin(\beta - \ln\lambda_i)\zeta_{max}}{(\beta - \ln\lambda_i)} \qquad (4.43)$$

where we have taken the truncation of the integral (4.38) (p.75) explicitly into account. Pizer et al. (1969) use (4.43) in the following iterative scheme to reduce the 'error ripples' associated with the sidelobes of each term in (4.43). $Y(e^{-\beta})$ is observed, and the number of peaks N is determined. The term in (4.43) associated with the highest peak is subtracted from $Y(e^{-\beta})$. The next highest peak is detected, and the corresponding term in (4.43) is also subtracted from $Y(e^{-\beta})$, and so on until N terms have been

eliminated. A new approximation to the highest peak is then obtained by subtracting from the original $Y(e^{-\beta})$, the terms in (4.43) associated with the other N-1 peaks. Then the terms corresponding to the highest peak and the lowest N-2 peaks are subtracted out, giving a new approximation to the second highest peak, and so on for the N peaks. The procedure is repeated until the peak positions no longer change significantly. The effect of this procedure is shown in Figs 4.5 and 4.6 which are reproduced from the article by Pizer et al. (1969). They used the test data

$$y(t) = e^{-.0001t} + 3e^{-.0003t} + 15e^{-.0015t}, \qquad (4.44)$$

and computed the respective exponents .000098, .000299, and .00155 without their subtraction procedure (fig. 4.5), and exponents .000099, .000301, and .00150, with error ripple subtraction (Fig. 4.6). Pizer et al. (1969) used $\zeta_{max} = 7.0$ for the results in Figs 4.5 and 4.6.

$R(\zeta)$ diverges primarily because, in a practical situation, the data $y(t)$ is available only for times in the range $0 \leq t \leq T_y$. Figures 4.7 and 4.8 show $R(\zeta)$ and $Y(e^{-\beta})$ for the single exponential example (4.40) with a restricted time range $0 \leq t \leq 4$. The divergence of $R(\zeta)$ in Fig. 4.7 occurs for lower ζ than it does in Fig. 4.1 (p.76). Fig. 4.8 shows the unsatisfactory $Y(e^{-\beta})$ obtained if ζ_{max} is set to the extremities of $R(\zeta)$ in Fig. 4.7. Pizer et al. (1969) and Gardner et al. (1959) extrapolate the data with a single exponential to extend the useful range of $R(\zeta)$. The latter authors conclude that, even though data extrapolation can introduce additional, unwanted, exponential components, the results are generally improved. However, extrapolation with

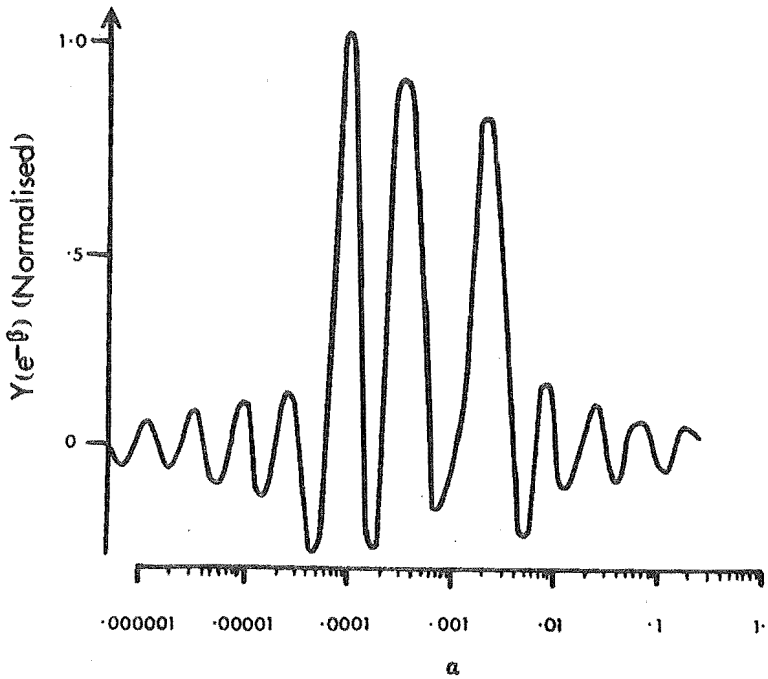


Fig. 4.5 $Y(e^{-\beta})$ as computed from the data of equation (4.44) by Pizer et al. (1969).

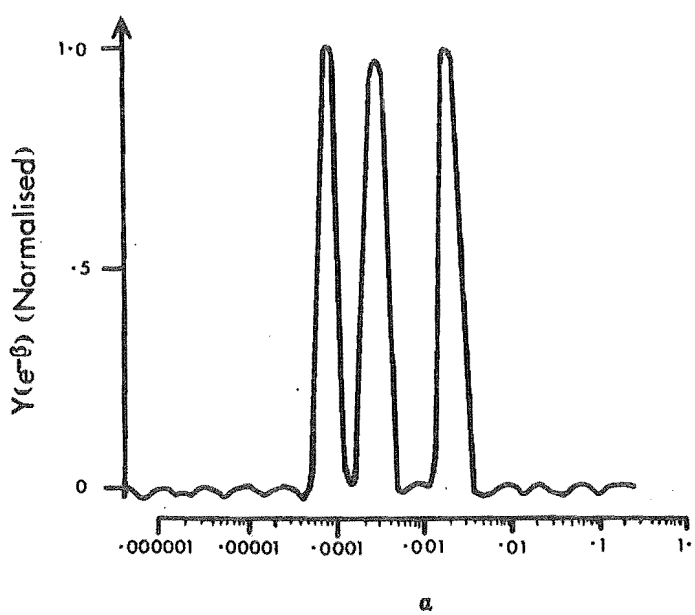


Fig. 4.6 $Y(e^{-\beta})$ computed by Pizer et al. (1969) after applying their error ripple subtraction procedure to the result of Fig. 4.5.

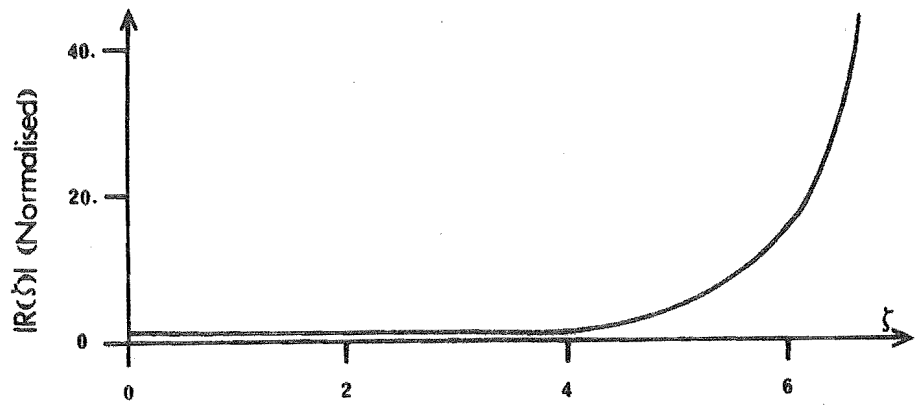


Fig. 4.7 $|R(\zeta)|$ computed from data given by equation (4.40) (p.75) but with $0 \leqslant t \leqslant 4$.

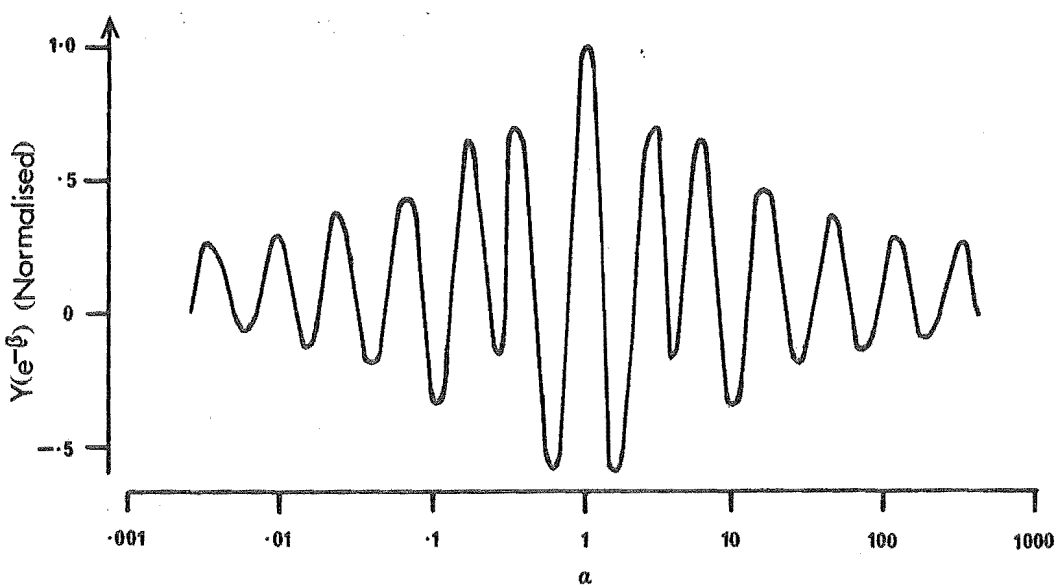


Fig. 4.8 $Y(e^{-\beta})$ computed from $R(\zeta)$ shown in Fig. 4.7.

an exponential function is bound to make the data appear more exponential in form, making it impossible for the method to check usefully that the data does actually represent an exponential sum.

The divergence of $R(\zeta)$ is also contributed to by error in the data. Gardner et al. (1959) present an example in which an error averaging approximately 3% (apart from two isolated data points in the tail of the curve with 16% and 122% respectively) is superimposed on the function

$$y(t) = 100\, e^{-.02t} \qquad 0 \leqslant t \leqslant 1100. \tag{4.45}$$

They find that, to reduce the extraneous peaks or sidelobes of $Y(e^{-\beta})$ to an acceptable lebel such as in Fig. 4.2 (p.76), it is necessary to have ζ_{max} approximately the same as if the data (4.45) were error free, but available for only $0 \leqslant t \leqslant 190$.

Before computing $f(1+j\zeta)$ using (4.37), Pizer et al. (1969) smooth and interpolate the data by taking its logarithm, and fitting a straight line to the three data points which lie on either side of the t value at which interpolation is required. They apply their smoothing and extrapolation procedures to the function

$$y(t) = e^{-.0001t} + 3e^{-.003t} + 15e^{-.0015t} \quad (0 \leqslant t \leqslant T_y) \tag{4.46}$$

with superimposed error of standard deviation 5%. They do not specify T_y, but state that the time domain is 'restricted'. Their estimates of the exponents .0001 and .0015 are in error by 6% and 100% respectively. This result is contrary to what we would expect. For a given T_y in (4.46), the term with exponent .0015 is more completely represented in the data than is the term with exponent .0001, and so the procedure should resolve it more accurately. However, Pizer et al. (1969) extrapolate with a single exponential fitted by eye to a plot of the tail portion of ln y(t). The behaviour of ln y(t) for large t is determined predominantly by the slowest

decaying term. Therefore, only the slowest decaying term is effectively extrapolated, thus allowing it to be estimated more correctly than those which decay faster.

Despite the preliminary smoothing and extrapolation of the data, the errors of 6% and 100% quoted above are no better than those obtained using other methods for estimating parameters of exponential functions (section 3.3.2).

4.3 ANALYTICAL CONTINUATION BY FUNCTIONAL REPRESENTATION

In section 4.2 we developed the theory, using integral transforms, for the analytic continuation of data $y(t)$ on to the Bromwich contour, so that $Y(\alpha)$ could be computed using the Laplace transform inversion formula (4.6) (p.68) The principal difficulty with the Mellin transform method of analytic continuation is shown in section 4.2.1 to be the diverging nature of $R(\zeta)$ when the data contains error or is truncated in time. In this section we attempt the analytic continuation of $y(t)$ by letting

$$y(t) = \sum_{m=1}^{M} Y_m \phi_m(t), \qquad (4.47)$$

and evaluating $y(p)$. The functions $\phi_m(t)$ are chosen so that $y(p)$ cannot diverge at the extremities of Br.

The ideal analytic continuation is obtained by letting

$$\phi_m(t) = e^{-\lambda_m t},$$

$$Y_m = A_m, \qquad (4.48)$$

and $$M = N,$$

but because the purpose of the analytic continuation is to estimate λ_m, A_m and N we must proceed otherwise.

Suppose the data y(t) consist of M samples at times t_n ($1 \leq n \leq M$) such that

$$t_M = T_y, \tag{4.49}$$

and

$$t_{n+1} > t_n \qquad \text{for all } n. \tag{4.50}$$

We let

$$y(t_n) = \sum_{m=1}^{M} Y_m e^{-\alpha_m t_n} \qquad \text{for } 1 \leq n \leq M, \tag{4.51}$$

where the α_m are specified real constants such that

$$\alpha_{m+1} > \alpha_m \qquad \text{for all } m. \tag{4.52}$$

From equation (4.51) the analytic continuation is

$$y(p) = \sum_{m=1}^{M} Y_m e^{-\alpha_m p}. \tag{4.53}$$

We restrict the region of analytic continuation to the right hand semicircle whose centre is the origin and whose radius is T_y (Fig. 4.9). $Y(\alpha)$ is then given by

$$Y(\alpha) = \sum_{m=1}^{M} Y_m \frac{\sin(\alpha-\alpha_m)T_y}{\pi(\alpha-\alpha_m)}. \tag{4.54}$$

There is no theoretical justification for restricting the region of validity of y(p) as in Fig. 4.9. If we allow y(p) to exist along the entire Bromwich contour, (4.54) reduces to a sum of delta functions, the position of which depend on only the α_m values and not on the data.

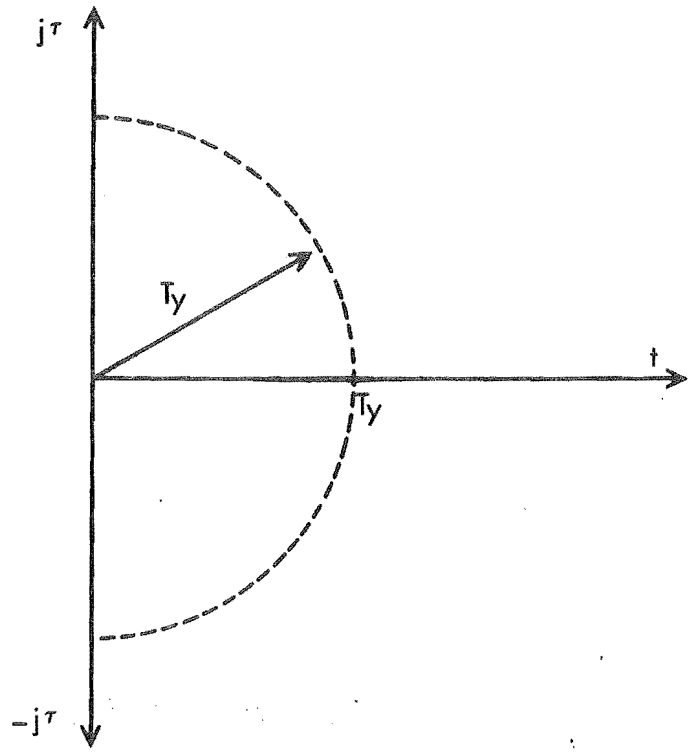


Fig. 4.9 Region of analytic continuation of y(t) in the complex p-plane.

Conversely, if the region of analytic continuation is restricted too severely, the peaks in (4.54) become wide, rendering them unable to faithfully represent the true $Y(\alpha)$. The region in Fig. 4.9 is intuitively chosen to lie between these extremes. The Y_m values are found by solving the M simultaneous equations

$$y(t_n) = \sum_{m=1}^{M} Y_m e^{-\alpha_m t_n} \qquad 1 \leq n \leq M. \tag{4.55}$$

The matrix inversion required to solve (4.55) is bound to be poorly conditioned (cf. Westlake, 1968) especially when M is large, because the terms $e^{-\alpha_m t_n}$ are large in the top row and

and left hand column of the matrix, falling away rapidly as m and n increase. However, even though the Y_m values obtained may depend significantly on M, we assume that the $Y(\alpha)$ computed from (4.54) does not show the same sensitivity.

In Figure 4.10 we show $Y(\alpha)$ computed from

$$y(t) = e^{-t} \qquad 0 \leqslant t \leqslant 4 \tag{4.56}$$

with

$$\alpha_m = .25(m-1) \tag{4.57}$$

and M set to 4, 7 and 12. In each case the Y_m values are sufficiently accurate to reconstruct the data $y(t_n)$ correct to seven significant figures when substituted back into equation (4.55). Although there is some deviation from the expected symmetry of the peak, the plots of $Y(\alpha)$ for M = 4 and 7 do suggest the existence of an exponent near $\alpha = 1$. The result for M = 12 is unacceptable in view of the ideal data used. We would prefer our procedure to improve as M increases, because y(t) is then more accurately defined. Figure 4.11 shows $Y(\alpha)$ computed from data generated using (4.56) but with

$$\alpha_m = .5 + (m-1)/(M-1) \tag{4.58}$$

for M = 12. Despite a similarly accurate matrix inversion to that obtained with α_m given by (4.57), the $Y(\alpha)$ in Fig. 4.11 does not correctly indicate the value of the exponent.

The results of Figs 4.10 and 4.11 are typical of many obtained, in that α_m values could not be specified to give a reliable $Y(\alpha)$, even with perfect data. Our assumption that $Y(\alpha)$ is not very sensitive to the Y_m values is not correct, and so we conclude that representation of data by (4.51) does not provide a useful analytic continuation.

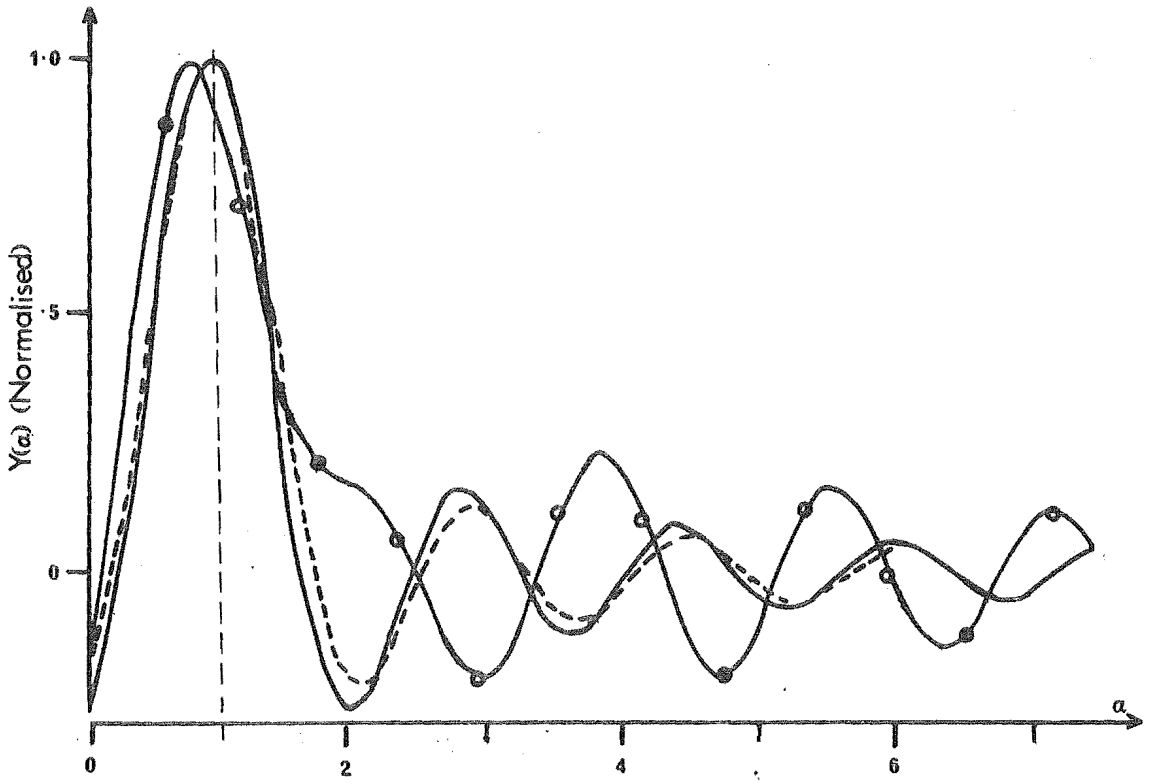


Fig. 4.10 $Y(\alpha)$ computed from the ideal data (4.56) with $M = 4$ (———), $M = 7$ (------), and $M = 12$ (o——o). $\alpha_m = .25(m-1)$.

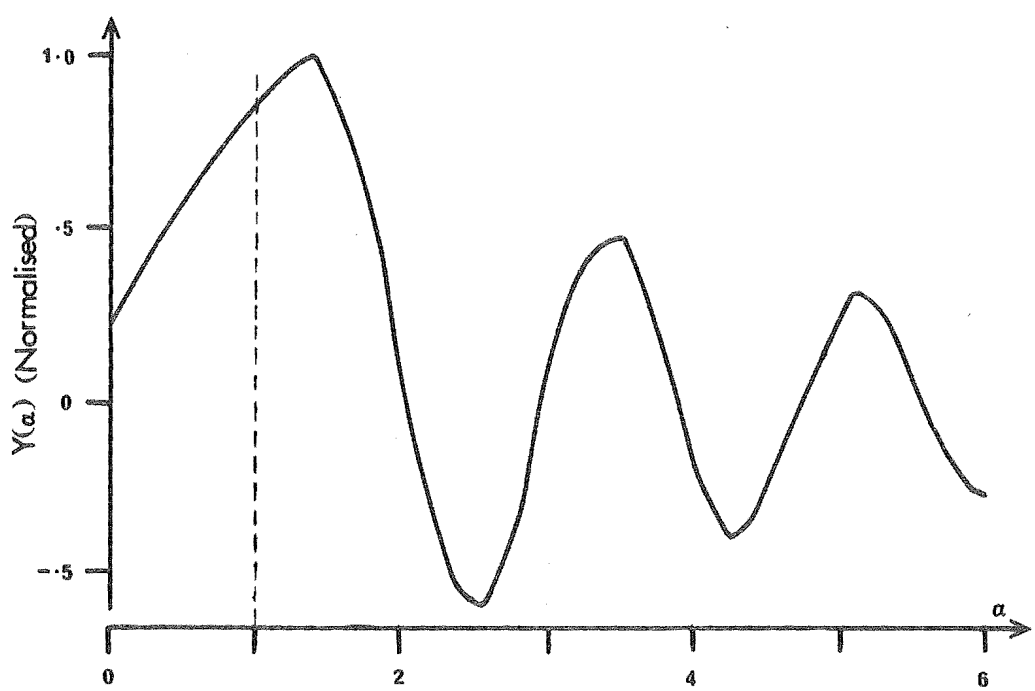


Fig. 4.11 $Y(\alpha)$ computed from the ideal data (4.56) with $M = 12$ and $\alpha_m = .5 + (m-1)/(M-1)$.

There are other possible approaches to analytic continuation using functional representation. For example one may let

$$\phi_m(t) = e^{-\sqrt{at}} L_m(bt), \tag{4.59}$$

where a and b are real positive constants and $L_m(x)$ is the Laguerre polynomial of order m (cf. Abramowitz and Stegun, 1965, p.775). The $\phi_m(t)$ in (4.59) has the desirable property that y(p) converges along both the real line and the Bromwich contour, and also the Y_m values are conveniently determined using the orthogonality property of Laguerre polynomials over the interval $0 \leq x \leq \infty$ (Abramowitz and Stegun, 1965, p.775).

However, the weight of evidence in sections 3.3, 4.2 and 4.3 suggests that the highly non-orthogonal nature of decaying exponential functions on the real line is such that any new approach is unlikely to improve the quality of the solutions obtained. Therefore, we do not pursue (4.59) further.

CHAPTER 5

CONSISTENCY CONDITIONS AND DECONVOLUTION

In section 3.4.1 (p.62) we mentioned a method of deconvolution in which the integral (3.56) (p.61) is approximated by a summation. We use this procedure to show that information as to the consistency of the data, is available in many deconvolution problems.

We define $f(x)$, $g(x)$ and $h(x)$ to exist only for x values in the intervals

$$0 \leqslant x \leqslant L_f, \tag{5.1}$$

$$0 \leqslant x \leqslant L_g, \tag{5.2}$$

and $$0 \leqslant x \leqslant L_h \tag{5.3}$$

respectively. That is, $f(x)$, $g(x)$ and $h(x)$ are of 'finite extent'. We assume that $f(x)$ is known at sample points $x = 0, T, 2T, \ldots, N_fT$ such that

$$(N_f+1)T = L_f, \tag{5.4}$$

and that N_g+1 values of $g(x)$ are known at the same sample spacing T. Inspection of (3.56) (p.61) yields

$$L_f = L_g + L_h - T \tag{5.5}$$

and

$$N_f = N_g + N_h. \tag{5.6}$$

For convenience we take

$$L_h = L_g \quad (\text{i.e. } N_h = N_g). \tag{5.7}$$

Assuming that f(x) changes smoothly between samples, the convolution integral (3.56) converts to the sums

$$f(nT) = T \sum_{m=0}^{n} g(nT-mT)h(mT) \quad \text{for } 0 \leq n \leq N_g, \tag{5.8}$$

and

$$f(nT) = T \sum_{m=n-N_g}^{N_g} g(nT-mT)h(mT) \quad \text{for } N_g \leq n \leq 2N_g. \tag{5.9}$$

Because the samples of f(x) and g(x) are known, (5.8) provides N_g+1 equations from which the N_g+1, unknown, h(x) samples may be determined. Therefore, the additional N_g+1 equations available from (5.9) are redundant. The number of redundant equations increases as the ratio L_g/L_h. When the functions f(x) and g(x) represent measurements, the redundant equations (5.9) can be used as a consistency check. (They are bound to be inconsistent.) Better still, all the equations implied by (5.8) and (5.9) can be solved together, giving a least squares solution for the samples of h(x) (cf. Patterson, 1950). There is, however, no systematic way of altering either h(x) or g(x) to obtain consistency when the accuracy of one of them is suspect.

To investigate how the redundancy implied by (5.8) and (5.9) appears when the deconvolution problem is expressed in the Fourier domain, we introduce the complex variables

$$w = u + jv, \tag{5.10}$$

and

$$z = x + jy. \tag{5.11}$$

The Fourier formulae (3.65) and (3.66) (p.64) can be rewritten as

$$F(w) = \int_{-\infty}^{\infty} f(x)\ e^{j2\pi wx}\ dx, \qquad (5.12)$$

and

$$f(z) = \int_{-\infty}^{\infty} F(u)\ e^{-j2\pi zu}\ du. \qquad (5.13)$$

We note that in (5.12) and (5.13) the integration is with respect to the real variables x and u respectively, so that the transforms may be applied directly to measured data. Equation (3.67) (p.65) which is given in terms of u, may be rewritten as

$$F(w) = G(w)\ H(w). \qquad (5.14)$$

Before developing the consistency conditions for deconvolution in the Fourier domain in section 5.2, we consider the use of (5.14) for deconvolution by inverse filtering.

5.1 INVERSE FILTERING IN THE COMPLEX PLANE

The standard inverse filtering approach to deconvolution is described in section 3.4.1 (p.65). It corresponds to using (5.14) with w replaced by the real variable u. Difficulties occur when G(u) is zero for some u, or G(u) decreases faster than F(u) as u becomes large. We noted in section 3.4.1 (p.65) that the first problem is solved, in practice, by rearranging the sampling of G(u) to avoid any points at which G(u) is zero. The second problem is dealt with by applying a filter function to the diverging F(u)/G(u). However, relevant information about the true H(u) is inevitably lost by this process, resulting in a distorted h(x). We consider how to reduce the information lost, by attempting inverse filtering for complex values of w.

Figure 5.1 is a normalised graph of $|H(w)|$ obtained by substituting

$$h(x) = e^{-2\pi x} \qquad 0 \leqslant x \leqslant 1 \qquad (5.15)$$

into (5.12) with F(w) and f(x) replaced by H(w) and h(x) respectively. The curves are plotted with respect to the real variable u, for v = 0 and v = -1.

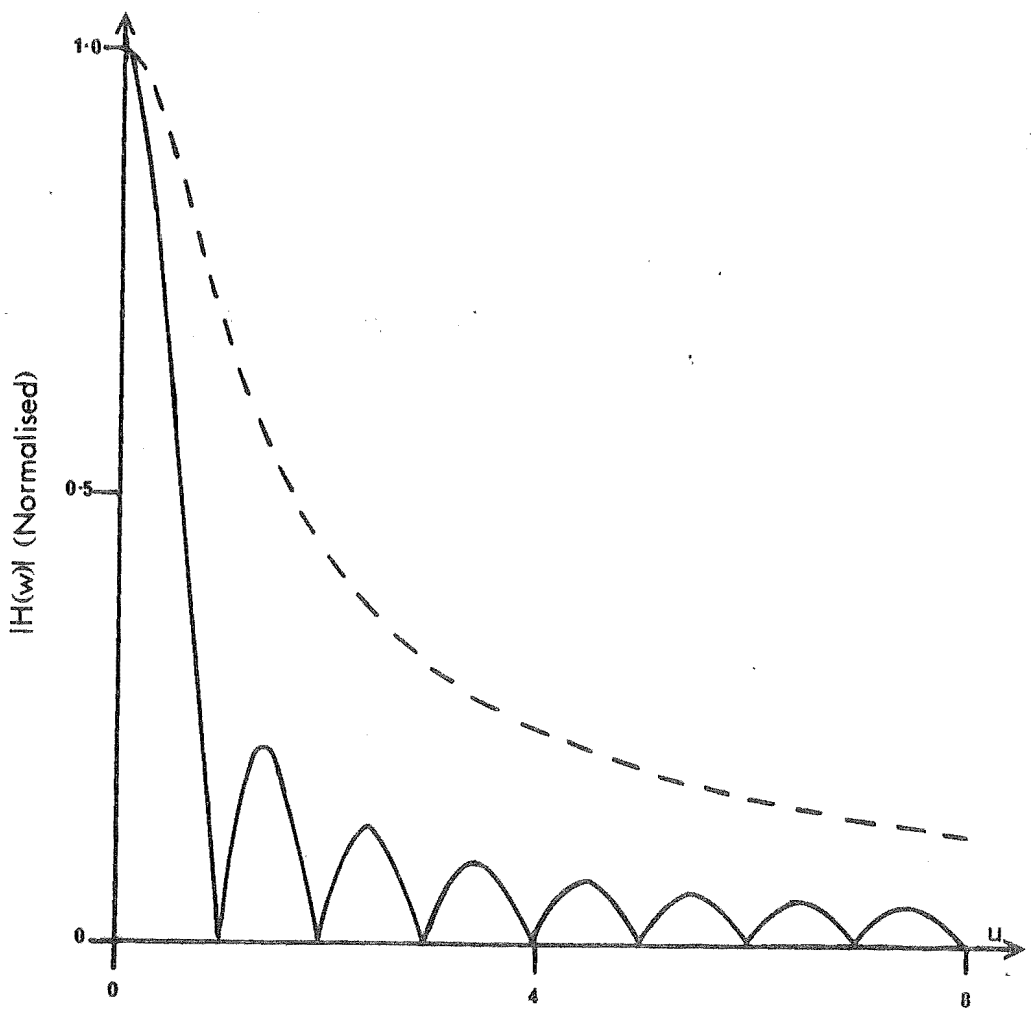


Fig. 5.1 $|H(w)|$ computed directly from h(x) in (5.15).
——— v = -1; ----- v = 0.

Clearly $|H(w)|$ converges faster when v = -1. In practice, H(w) computed from F(w)/G(w) contains noise whose behaviour we cannot predict as v varies. But if, for example, the noise

requires the filtering of H(w) at the same value of u, irrespective of v, then it is clear from Fig. 5.1 that we lose less information about the true H(w) by choosing v = -1.

We present computational examples in which h(x), given by (5.15), is analytically convolved with

$$\begin{aligned} g(x) &= 1 && \text{for } 0 \leqslant x \leqslant 1 \\ &= 0 && \text{otherwise,} \end{aligned} \qquad (5.16)$$

to give

$$\begin{aligned} f(x) &= \frac{1}{2\pi}\left(1 - e^{-2\pi x}\right) && \text{for } 0 \leqslant x \leqslant 1 \\ &= \frac{1}{2\pi}\left(e^{-2\pi(x-1)} - e^{-2\pi}\right) && \text{for } 1 \leqslant x \leqslant 2 \\ &= 0 && \text{otherwise.} \end{aligned} \qquad (5.17)$$

The Fourier transforms F(w) and G(w) are computed as indicated in section 8.1 (p.156). We only compare results for v = 0 and v = -1 because v = 0 corresponds to the standard inverse filtering deconvolution procedure (3.68) (p.65) and v = -1 is the contour along which H(w), theoretically, converges most rapidly when h(x) is given by (5.15).

The details of the tests carried out are given in Table 5.2. The noise superimposed on f(x) is normally distributed with standard deviation a given fraction of the value of f(x). We call H(u) and H(u-j) the transforms calculated from (5.14) when v = 0 and -1 respectively. Because H(u) and H(u-j) tend to diverge as u becomes large, filters (or 'windows') are applied to them as given in Table 5.2. The cosine bell window of width u_d, smoothly reduces the function to zero for $|u| > u_{max}$, whereas the rectangular window cuts it off sharply at $|u| = u_{max}$ (section 8.1, p.155). h(x) is calculated from

Table 5.2

Details of Inverse Filtering Tests

Standard Deviation of Noise on f(x)	Figure showing $\|H(u)\|$	Figure showing $\|H(u-j)\|$	Window applied to H(u)	Window applied to H(u-j)	Figure showing $h_0(x)$	Figure showing $h_{-1}(x)$
0	5.3	5.3	cosine bell $u_{max} = 12.9$ $u_d = 3.0$	rectangular $u_{max} = 7.0$	5.4	5.5
.05	5.6	5.6	cosine bell $u_{max} = 4.1$ $u_d = 2.6$	rectangular $u_{max} = 3.0$	5.7	5.8

H(w) using (5.13), with f(z) and F(u) replaced by h(x) and H(w) respectively (section 8.1, p.157). We denote by $h_0(x)$ the h(x) corresponding to H(u) and by $h_{-1}(x)$ the h(x) corresponding to H(u-j). Table 5.2 also lists the Figures which show the actual computed samples of the various functions.

We note that even with no noise added to f(x), both $|H(u)|$ and $|H(u-j)|$ diverge as u becomes large (Fig. 5.3). This is due to the necessarily approximate Fourier transform calculations (section 8.1, p.154). The consequences are evident in $h_0(x)$ (Fig. 5.4) and $h_{-1}(x)$ (Fig. 5.5). We note that in the latter case, large oscillations occur for $x < 0$. In section 8.1 (p.157) it is shown that to calculate $h_{-1}(x)$ from H(u-j), we first compute an inverse transform using (5.13) as if we actually have H(u), and then multiply the result by $e^{-2\pi x}$. The oscillations in Fig. 5.5 are the inaccuracies of the inverse transform multiplied by $e^{-2\pi x}$. However, we have specified the region of existence of f(x) by (5.17) and g(x) by (5.16), and so, by inspection of the convolution integral (3.56) (p.61), h(x) exists only for $0 \leqslant x \leqslant 1$. We therefore ignore its behaviour outside $0 \leqslant x \leqslant 1$, and observe that the results $h_0(x)$ (Fig. 5.4) and $h_{-1}(x)$ (Fig. 5.5) are of similar quality.

With 5% noise added to f(x), $|H(u)|$ diverges markedly in contrast to $|H(u-j)|$ (Fig. 5.6). Filtering of H(u) as indicated in Table 5.2 gives the surprisingly good reconstruction $h_0(x)$ in Fig. 5.7. $h_{-1}(x)$ (Fig. 5.8) is a good approximation to h(x) for $x > .1$, but it rounds off the initial sharp peak evident in the true h(x).

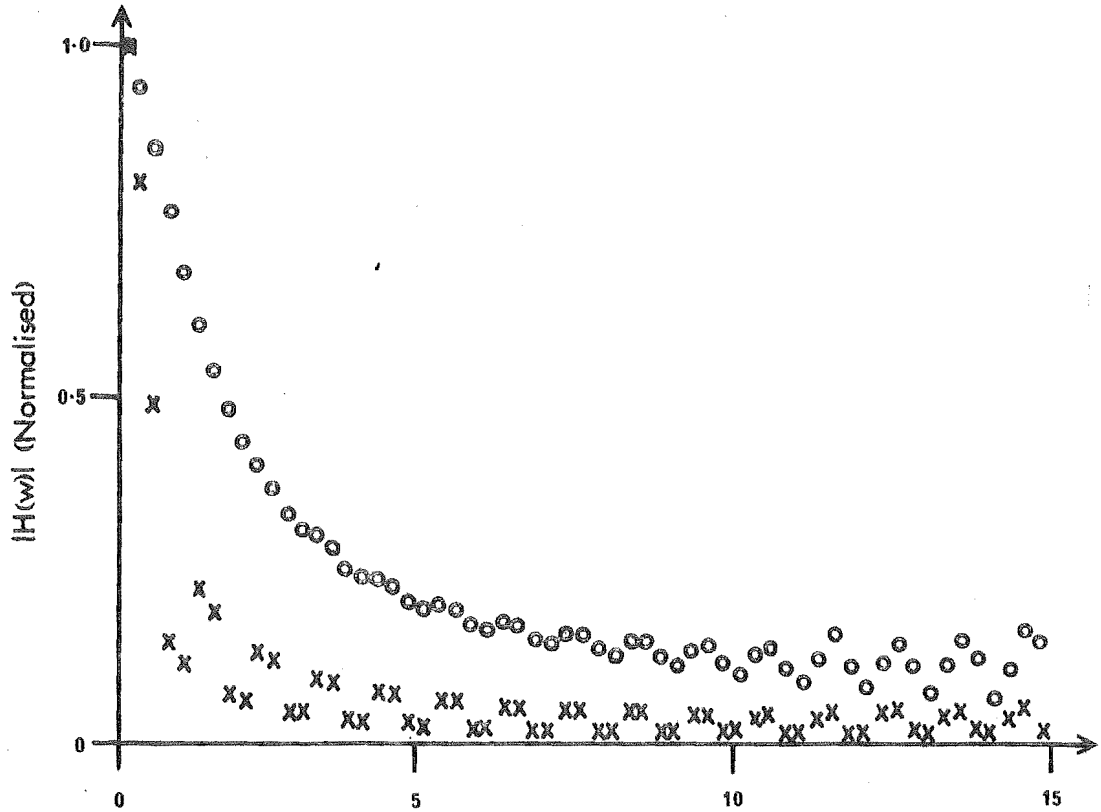


Fig. 5.3 |H(w)| when f(x) contains no error. OOOO v = 0; XXXX v = -1.

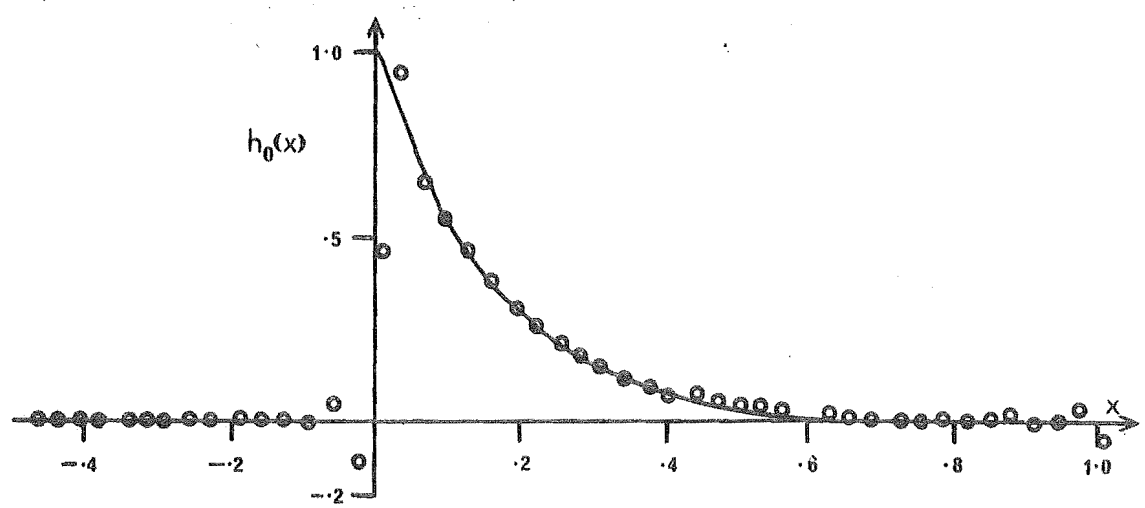


Fig. 5.4 $h_0(x)$ reconstructed from H(u) shown in Fig. 5.3. ——— theoretical h(x); OOO reconstruction.

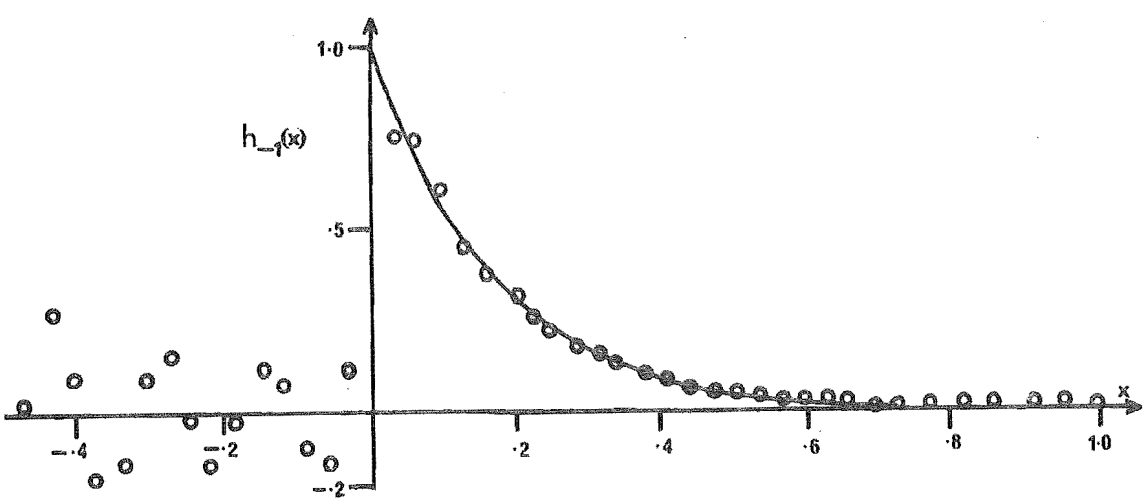


Fig. 5.5 $h_{-1}(x)$ reconstructed from H(u-j) shown in Fig. 5.3. ——— theoretical h(x); OOO reconstruction.

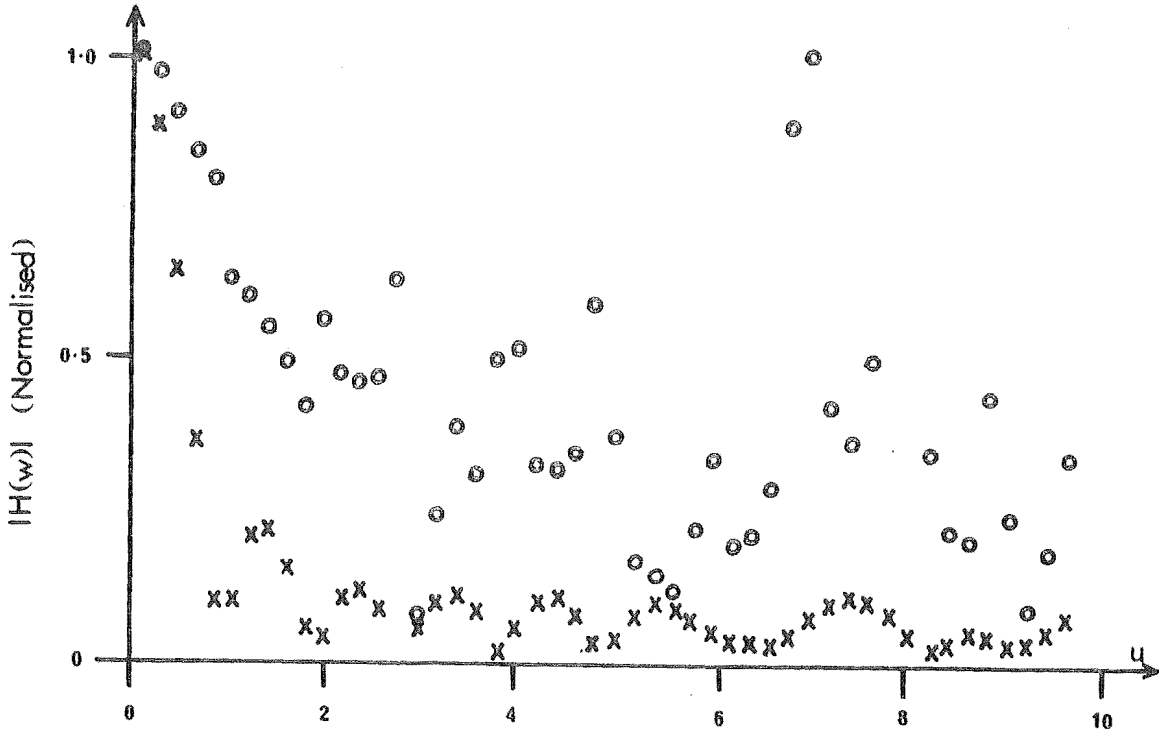


Fig. 5.6 $|H(w)|$ when f(x) contains 5% error. OOOO v = 0; XXXX v = -1.

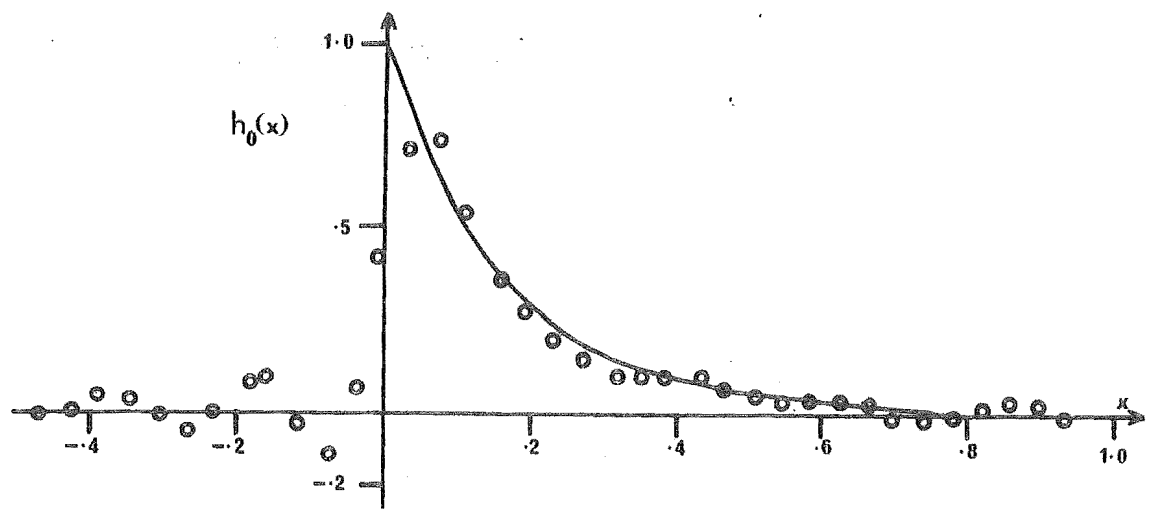


Fig. 5.7 $h_0(x)$ reconstructed from H(u) shown in Fig. 5.6. —— theoretical h(x); OOO reconstruction.

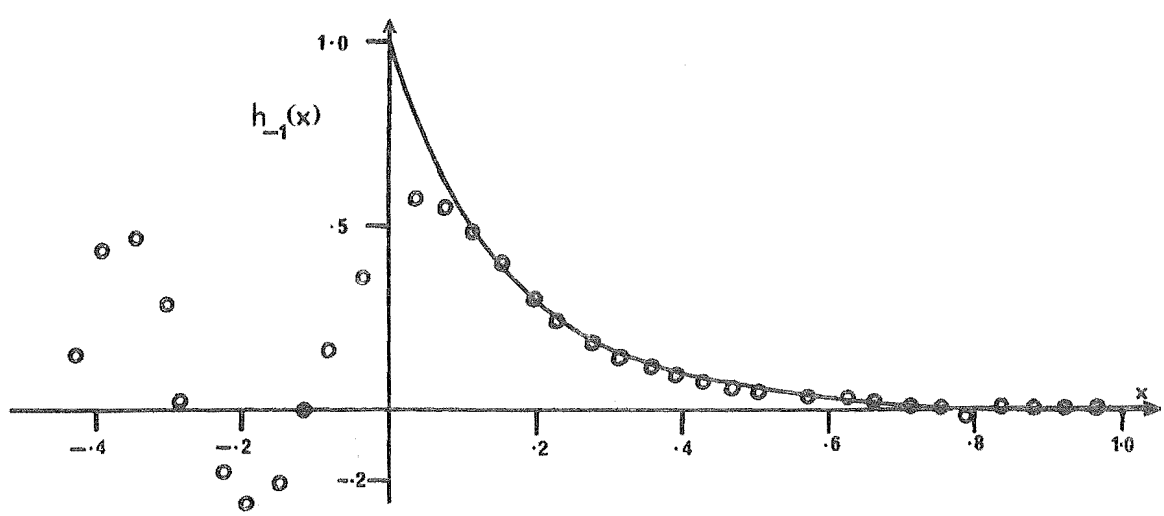


Fig. 5.8 $h_{-1}(x)$ reconstructed from H(u-j) shown in Fig. 5.6. —— theoretical; OOO reconstruction.

The above tests show that there can be some advantage in carrying out deconvolution by inverse filtering in the complex plane with v chosen so that the true H(w) converges fastest. However, in a practical deconvolution problem, the behaviour of the true H(w) is unknown (otherwise there is no need to deconvolve), and so it is impossible to specify the optimum v a priori. Further evaluation of the method is required to assess the effects of various filters applied to H(w), to test various forms for g(x) and h(x), and to see if the best value of v can be found by trial.

5.2 THE ZEROS OF THE FOURIER TRANSFORM

In the example at the beginning of chapter 5 (p.89), we assumed that f(x), g(x) and h(x) are of finite extent. Many observable quantities can be assumed either to have insignificant value outside a finite range, or to fall asymptotically to zero for large values of the independent variable. Therefore, we impose only a weak constraint on f(x) if we require that either

$$f(x) \text{ exists for } X_f \leq x \leq X_f + L_f, \tag{5.18}$$

or

$$\lim_{|x|\to\infty} f(x) = A_f\, e^{-x^2/a_f^2}, \tag{5.19}$$

where A_f and a_f are positive real constants. When either (5.18) or (5.19) holds, F(w) is an entire or integral function of exponential type given by (cf. Paley and Wiener, 1934)

$$F(w) = b\, e^{-\psi(w)} \prod_m \left(1 - \frac{w}{w_{F,m}}\right) \tag{5.20}$$

In equation (5.20), $\psi(w)$ is a polynomial in w, b is a real

constant, and the $w_{F,m}$ are the zeros of F(w), of which there may be an infinite number (Paley and Wiener, 1934). If (5.18) holds, $\psi(w)$ is of order 1, whereas it is of order 2 if (5.19) holds. The property of entire functions which we exploit is that they are characterised by the positions of their zeros $w_{F,m}$. That is, the values of w for which F(w) = 0.

Because we assume f(x) to be real (the convolution integral (3.56) (p.61) and Fourier transform (5.12) and (5.13) can be applied for f(x) complex, however), F(w) given by (5.12) (p.91) is conjugate symmetric, which means that

$$F(-w^*) = F^*(w), \tag{5.21}$$

where the asterisk denotes the complex conjugate.

Because of (5.21), the zeros $w_{F,m}$ of F(w) are distributed symmetrically about the imaginary (jv) axis. That is, if

$$w_{F,m} = u_{F,m} + jv_{F,m} \quad (u_{F,m} \neq 0;\ u_{F,m},\ v_{F,m} \text{ real}) \tag{5.22}$$

is a zero of F(w), then so is $-u_{F,m} + jv_{F,m}$. Individual zeros may, however, occur on the imaginary axis. Because the zeros either lie on the imaginary axis or are symmetrically placed about it, we show only zeros in the right half w-plane or on the imaginary (jv) axis in all diagrams indicating their positions. We denote by ζ_F the set of zeros of F(w), and by ξ_F the set of zeros of F(w) which have non-zero imaginary parts. That is

$$v_{F,m} \neq 0 \qquad \text{for all } w_{F,m} \text{ in } \xi_F. \tag{5.23}$$

ξ_F is a subset of ζ_F. We order ζ_F so that its first M_F members are the members of ξ_F, where M_F is the cardinal number (i.e. number of members) of ξ_F.

When (5.18) applies, ζ_F necessarily has an infinite number of distinct members (cf. Paley and Wiener, 1934). However, we shall see in section 7.1 that M_F can always be taken as finite when F(w) is computed from measured data. When (5.19) applies, ζ_F has only a finite number of members all of which also belong to ξ_F, because it is unlikely that any zero computed from measured data will lie exactly on the real line. A further consequence of (5.19) is that both f(z) and F(w) are entire functions of order 2. If f(x) is real, its zeros are distributed symmetrically about the real axis, and if f(x) is positive, any of its zeros on the real axis must be of even multiplicity. (A pair of zeros at the same point in the complex plane have multiplicity 2).

Bates (1969a, 1969b) and Bates and Napier (1972) use the zeros of F(w) when f(x) is of finite extent, to reduce ambiguity in the interpretation of interferograms, for which accurate phase measurements are difficult to obtain.

5.2.1 Consistency Conditions

If f(x) satisfies (5.18) or (5.19) we assume that g(x) and h(x) do also, so that either

$$g(x) \text{ exists for } X_g \leqslant x \leqslant X_g + L_g, \tag{5.24}$$

and

$$h(x) \text{ exists for } X_h \leqslant x \leqslant X_h + L_h, \tag{5.25}$$

or

$$\lim_{|x|\to\infty} g(x) = A_g e^{-x^2/a_g^2}, \tag{5.26}$$

and

$$\lim_{|x|\to\infty} h(x) = A_h e^{-x^2/a_h^2}, \tag{5.27}$$

where A_g, a_g, A_h and a_h are positive real constants.

Corresponding to ζ_F and ξ_F, we define the sets ζ_G and ξ_G of zeros of G(w), and the sets ζ_H and ξ_H of zeros of H(w).

Following immediately from (5.14) (p.91) we introduce the 'consistency condition'

$$\zeta_F = \zeta_G \cup \zeta_H \tag{5.28}$$

where $\cup$ denotes union (or merging). Equation (5.28) states that the zeros of H(w) and G(w) must also be the zeros of F(w).

In Fig. 5.9 the circles denote some of the zeros of G(w) when g(x) is given by equation (5.16) (p.93), and the crosses denote some of the zeros of F(w) with f(x) given by (5.17). The ideal nature of this example means that there is actually an infinite number of zeros of G(w) and of F(w). We note that each zero of ζ_G is coincident with a zero of ζ_F as required by (5.28). The members of ζ_F which do not correspond with any member of ζ_G, constitute the set ζ_H.

We consider the implications of (5.28) in sections 5.2.2, and exploit them in chapters 6 and 7.

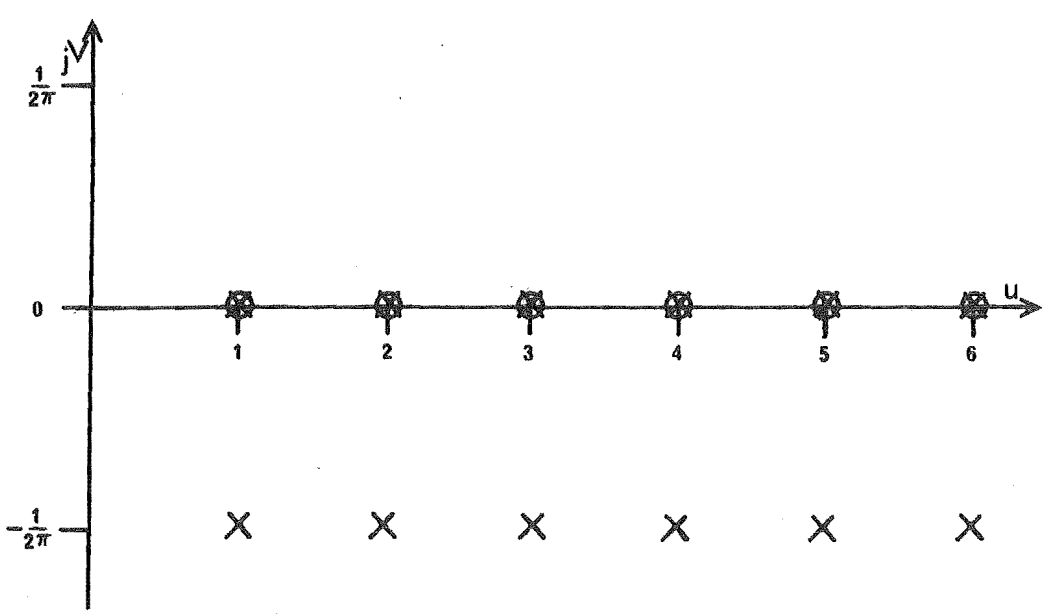


Fig. 5.9 Zeros of the Fourier transforms of g(x) (5.16) and f(x) (5.17) (p.93). OOO zeros of G(w); XXX zeros of F(w).

5.2.2 Consequences of Inconsistent Data

Any experimentally determined quantity contains error, and we suggested at the beginning of chapter 5, that a consistency check is therefore bound to fail. When ζ_F and ζ_G are computed from actual experimental data we do not expect (5.28) to hold, as no zero $w_{G,m}$ is likely to correspond exactly with a zero $w_{F,m}$. We are concerned with the degree of correspondence between the zeros of F(w) and those of G(w), and with what can be done to make the data consistent.

We denote by $r_m e^{j\theta_m}$ the distance in the complex w-plane between a zero $w_{F,m}$ and its nearest $w_{G,m}$. We consider the members of ζ_F to be ordered so that we have the best correspondence between zeros of F(w) and G(w), in the sense that

$$R = \sum_{m=1}^{M_G} r_m \qquad (5.30)$$

is minimised, where M_G is the cardinal number of ζ_G. Therefore,

$$r_m e^{j\theta_m} = w_{F,m} - w_{G,m}. \qquad (5.31)$$

We can do nothing to make data consistent without further information, but it is possible that an experimenter will know that the errors either in f(x) or in g(x) are significantly the greater.

If f(x) is more accurate, we replace the zeros ζ_G by the zeros ζ_{GF}, which are the members of ζ_F closest to the members of ζ_G in the sense defined by (5.30). This operation gives a new G(w) (written $G_F(w)$) such that

$$G_F(w) = G(w) \prod_{m=1}^{M_G} \left(\frac{w-w_{F,m}}{w-w_{G,m}}\right), \qquad (5.32)$$

$G_F(w)$ is consistent with F(w) in the sense defined by (5.28).

If g(x) is the more accurate, we form a new set ζ_{FG} which consists of ζ_F with the first M_G members replaced by the corresponding zeros $w_{G,m}$. This gives a new F(w) (written $F_G(w)$) such that

$$F_G(w) = F(w) \prod_{m=1}^{M_G} \left(\frac{w-w_{G,m}}{w-w_{F,m}}\right). \qquad (5.33)$$

However, (5.33) will not, in general, satisfactorily correct f(x), because it is unlikely that only the first M_G zeros in ζ_F will be in error, while the remaining M_F-M_G are correct. Therefore, the shifting process of (5.33) should be followed by a procedure in which the remaining unaltered zeros of ζ_F are adjusted to give some sort of 'best fit' to the original f(x).

It is important to note that the zero shifting operations (5.32) and (5.33) affect the whole of g(x) and f(x). Thus equation (5.28) is an advance on equations (5.8) and (5.9) (p.90) because the latter, although providing consistency information, do not suggest a procedure for correcting the inconsistent data.

It is difficult to make any general statement as to the effect on g(x) of the zero shifting in (5.32), as it depends on g(x) itself, on the positions of the zeros being shifted, and on the distance ($r_m e^{j\theta_m}$) that they are moved. Bates (1965) shows that shifting the zeros of G(w) when g(x) is of finite extent, does not alter the width of g(x).

CHAPTER 6

DECONVOLUTION OF COMPLETE DATA

The consistency condition for deconvolution which we developed in section 5.2.1 (p.101), is based on the relationship (5.14) (p.91) between the Fourier transforms of the quantities in the convolution integral (3.56) (p.61). We note from equation (5.12) (p.91), that to calculate the Fourier transform F(w) it is necessary, in principle, to know the function f(x) throughout the interval $-\infty < x < \infty$. Even though we may not know f(x) over the entire interval $-\infty < x < \infty$, we say it is 'complete', if the resulting error in F(w) is small. The deconvolution of complete data is a Type A problem in the review in section (3.4.1).

Any experimentally determined quantity contains error, and we acknowledged in section 5.2.2 that F(w) will not be zero at any of the values of w for which G(w) is zero. That is, none of the zeros of F(w) will exactly correspond with any of the zeros of G(w). However, we realised that if the zeros of F(w) and G(w) are close enough to enable matching pairs to be identified, we can shift them to obtain consistency. In this chapter we attempt to apply these ideas to data from X-ray diffraction experiments.

6.1 CALCULATION OF ZEROS USING HERMITE FUNCTIONS

The data shown in Figs 6.1 and 6.2 are typical of the f(x) and g(x) from X-ray diffraction measurements (section 6.3). We note that both f(x) and g(x) fall asymptotically to zero as $|x|$ becomes large, and we consider that the requirement (5.19) (p.98) is satisfied, thus making F(w) and G(w) entire

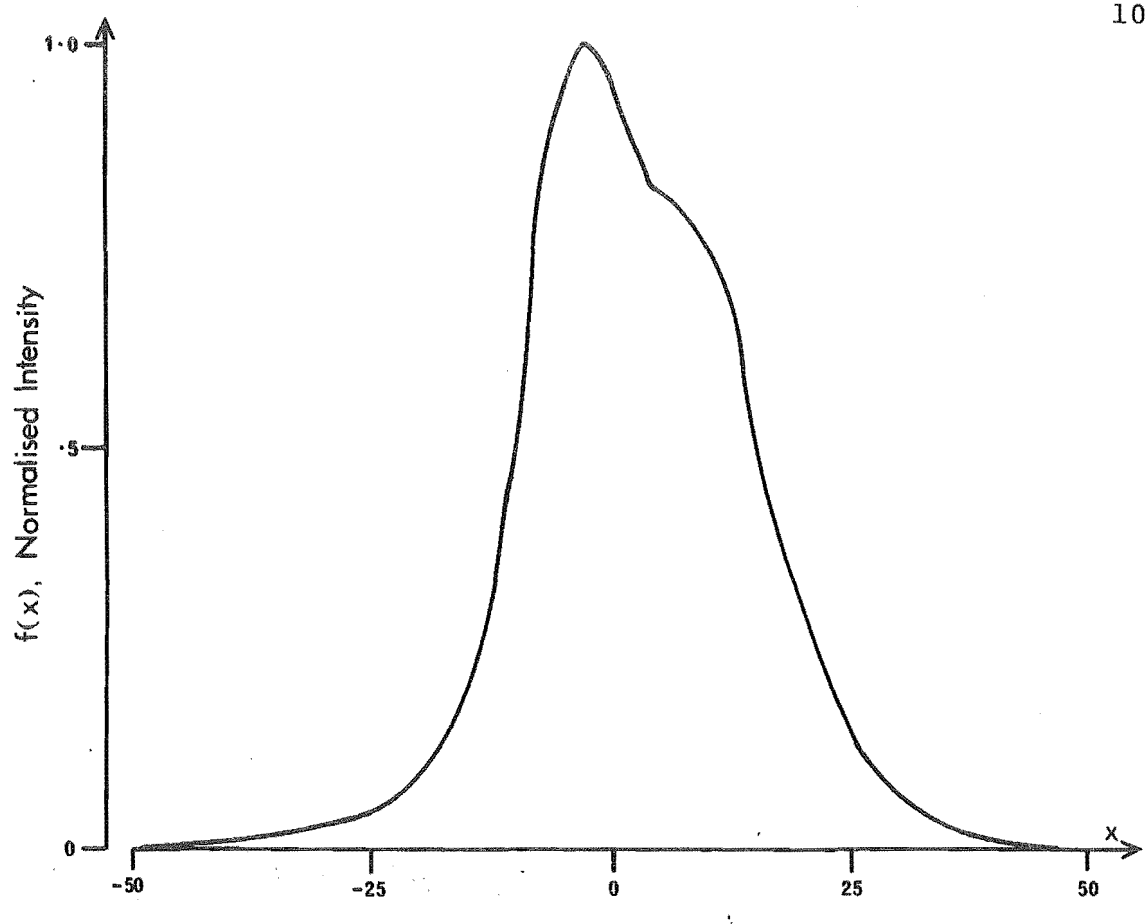


Fig. 6.1 A typical f(x) from X-ray diffraction experiments.

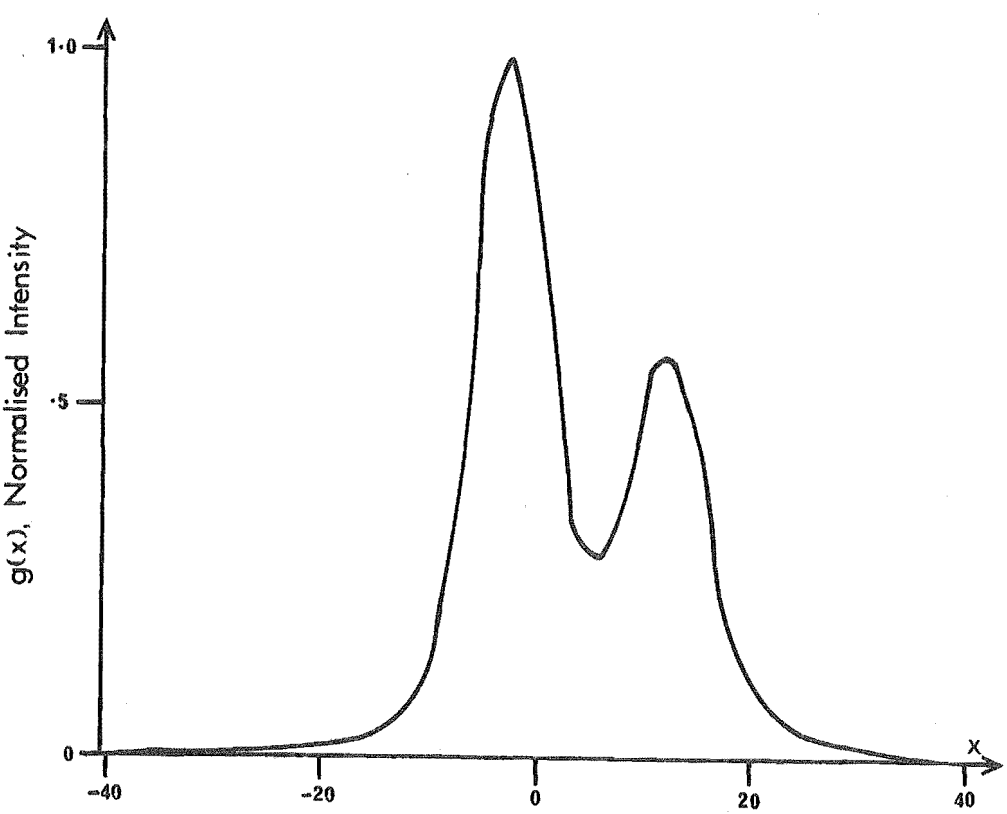


Fig. 6.2 A typical g(x) from X-ray diffraction experiments.

functions of order 2.

A convenient representation for f(x) satisfying (5.19), is in terms of Hermite functions. Accordingly we let

$$f(x) = e^{-x^2/a_f^2} \sum_{m=0}^{M_F} F_m \, He_m(2x/a_f), \tag{6.1}$$

where $He_m(y)$ is the Hermite polynomial of order m (cf. Abramowitz and Stegun, 1965, p.775). The estimation of the positive or zero integer M_F, and the real constants a_f and F_m is discussed in sections 6.1.1 and 6.1.2.

The Fourier transform of f(x) is given by (cf. Campbell and Foster, 1961)

$$F(w)=a_f\sqrt{\pi}\, e^{-\pi^2 a_f^2 w^2} \sum_{m=0}^{M_F} j^m F_m He_m(2\pi a_f w) \tag{6.2}$$

which is of the form

$$F(w) = A_F \, e^{-\pi^2 a_f^2 w^2} \, P_{M_F}(w). \tag{6.3}$$

In equation (6.3) A_F is given by

$$A_F = a_f \sqrt{\pi} \, (j2\pi a_f)^{M_F} \, F_{M_F}, \tag{6.4}$$

and $P_{M_F}(w)$ is a polynomial of order M_F given by

$$P_{M_F}(w) = a_f\sqrt{\pi}/A_F \sum_{m=0}^{M_F} j^m F_m He_m(2\pi a_f w) \tag{6.5}$$

or

$$P_{M_F}(w) = \prod_{m=1}^{M_F} (w - w_{F,m}). \tag{6.6}$$

The exponential term in (6.3) is never zero in the finite part of the complex w-plane, and so the zeros of F(w) are the M_F zeros $w_{F,m}$ of $P_{M_F}(w)$. As explained in section 5.2 (p.99) the condition that f(x) is real ensures that the zeros of F(w) lie either on the imaginary (jv) axis or symmetrically about it. Also, the existence of at least one zero on the imaginary axis means that f(x) takes on negative values for some x (Roman and Marathay, 1963). If M_F is odd, there must always be a zero of F(w) on the imaginary axis. Because we will be dealing with functions f(x), g(x) and h(x) which are positive, we therefore require M_F, M_G and M_H to be even.

To calculate the zeros ξ_F from data f(x) we proceed as follows:

Procedure P1 for calculating zeros of F(w)

P1.1 Determine the parameters a_f, M_F and F_m in equation (6.1) using the methods detailed in sections 6.1.1 and 6.1.2.

P1.2 Compute the M_F complex coefficients of the polynomial $P_{M_F}(w)$ using equation (6.5). (The precise formula for the coefficients is given in section 8.3.4, p.162).

P1.3 Find the zeros of $P_{M_F}(w)$ as described in section 8.5.

We will also want to reconstruct f(x) from the zeros ξ_F.

Procedure P2 for reconstructing f(x) from the zeros ξ_F.

P2.1 Using the formula given in section 8.3.4 (p.162) calculate the M_F+1 coefficients of the polynomial $Q_{M_F}(x)$ in

$$f(x) = e^{-x^2/a_f^2} Q_{M_F}(x) \qquad (6.7)$$

from the zeros ξ_F.

P2.2 Evaluate (6.7) at the required values of x.

Similar to the representation of f(x) in equation (6.1) we let

$$g(x) = e^{-x^2/a_g^2} \sum_{m=0}^{M_G} G_m He_m(2x/a_g), \tag{6.8}$$

and the unknown

$$h(x) = e^{-x^2/a_h^2} \sum_{m=0}^{M_H} H_m He_m(2x/a_h). \tag{6.9}$$

The quantities a_g, M_G, G_m, a_h, M_H and H_m in equations (6.8) and (6.9) correspond to the quantities a_f, M_F and F_m in the expansion (6.1). The Fourier transforms of g(x) and h(x) are, respectively,

$$G(w) = a_g\sqrt{\pi}\, e^{-\pi^2 a_g^2 w^2} \sum_{m=0}^{M_G} j^m G_m He_m(2\pi a_g w), \tag{6.10}$$

and

$$H(w) = a_h\sqrt{\pi}\, e^{-\pi^2 a_h^2 w^2} \sum_{m=0}^{M_H} j^m H_m He_m(2\pi a_h w). \tag{6.11}$$

We note that the relationship (5.14) (p.91) between F(w), G(w) and H(w) means that

$$a_f^2 = a_g^2 + a_h^2 \tag{6.12}$$

and

$$M_F = M_G + M_H. \tag{6.13}$$

Although the procedures P1 and P2 describe operations on f(x) and the zeros ξ_F, we will refer to them with respect to the functions g(x) and h(x) and the zeros ξ_G and ξ_H, making the appropriate changes in notation.

6.1.1 Representation of Data by Hermite Functions

The Hermite polynomials $He_m(2x/a_f)$ are orthogonal over the interval $-\infty < x < \infty$ with a weighting function e^{-2x^2/a_f^2}. This means that (cf. Abramowitz and Stegun, 1965, p.775),

$$\int_{-\infty}^{\infty} e^{-2x^2/a_f^2}\, He_n(2x/a_f)\, He_m(2x/a_f)\,dx = 0 \qquad m \neq n$$

$$= \nu_n\, a_f/2 \qquad m = n \tag{6.14}$$

where

$$\nu_n = \sqrt{2\pi}\, n! \tag{6.15}$$

Using the relation (6.14) we find from (6.1) (p.106) that

$$\int_{-\infty}^{\infty} e^{-x^2/a_f^2}\, He_n(2x/a_f)\, f(x)\, dx = F_n\, \nu_n\, a_f/2, \tag{6.16}$$

which provides a convenient method for calculating the F_m from data $f(x)$, once a_f has been determined.

Procedure P3 for calculating the coefficients in the expansion (6.1).

P3.1 Multiply the data $f(x)$ by $e^{-x^2/a_f^2}\, He_n(2x/a_f)$.

P3.2 Integrate the result as described in section 8.3.2 (p.159), and divide the integral by $\nu_n\, a_f/2$ to obtain F_n.

As in procedures P1 and P2, we will refer to P3 without qualification, to indicate finding the coefficients G_m in the Hermite function representation (6.8) of $g(x)$.

It is unlikely that any experimental data $f(x)$ is exactly representable by a finite number of terms in the Hermite expansion (6.1). For given a_f and M_F we define the root mean square (RMS) error by

$$E = \left(\sum_{i=1}^{N} (f(x_i) - f_r(x_i))^2/N \right)^{\frac{1}{2}}, \tag{6.17}$$

where there are N data points of $f(x)$ given at $x = x_1, x_2, \ldots, x_N$, and $f_r(x)$ is the value of the Hermite function representation (6.1) (p.106) with $f(x)$ replaced by $f_r(x)$.

The parameter a_f in equation (6.1) enables the x values to be scaled to correspond with the particular f(x) being represented. As a_f increases, the Gaussian weight function e^{-x^2/a_f^2} becomes wider. Therefore, we expect the value of a_f to affect both the accuracy of the representation of the data by the Hermite expansion (6.1), and the positions of the zeros of F(w). In Fig. 6.3 the zeros ξ_F are computed from the f(x) of Fig. 6.1 using procedures P3 and P1 with $M_F = 18$.

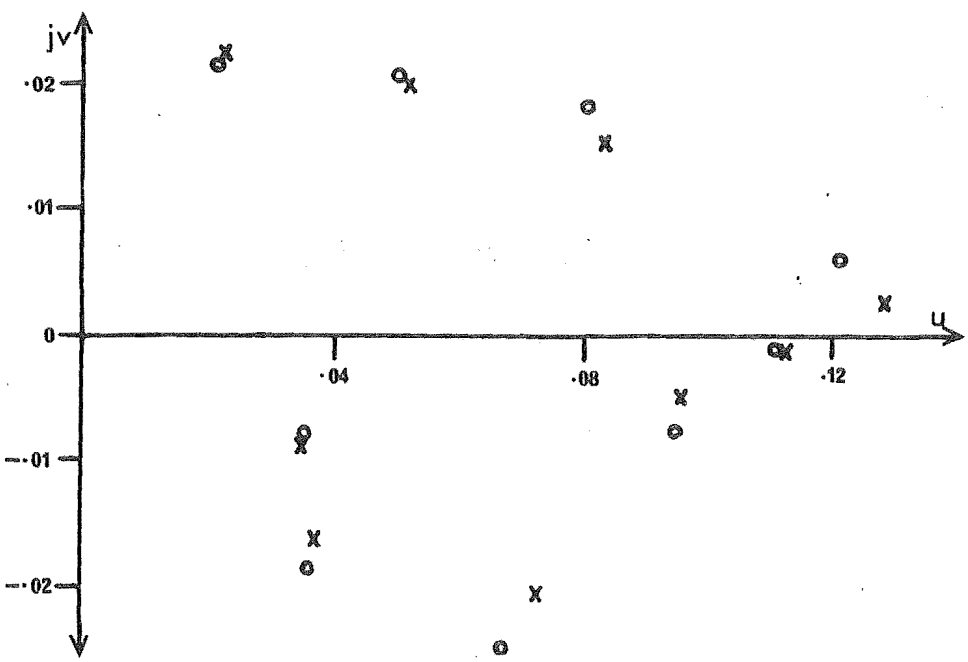


Fig. 6.3 Sensitivity of the zeros of F(w) to the parameter a_f. OOO $a_f = 9.66$; XXX $a_f = 9.18$

The crosses correspond to $a_f = 9.18$ and the circles correspond to $a_f = 9.66$. We see that the value of a_f affects the positions of the zeros, but has little affect on the mean squared error, which is 2.5 when $a_f = 9.18$ and 1.6 when $a_f = 9.66$.

There are various possible ways of estimating a_f (and a_g and a_h), which we describe in the procedures P4 - P8 which

follow. However, we first note that if f(x) is actually a Gaussian function (e^{-x^2/a_f^2}), then the representation (6.1) (p.106) requires only the single term for m = 0. For this case

$$\int_{-\infty}^{\infty} f(x)\ dx = F_0\ a_f\ \sqrt{\pi}, \tag{6.18}$$

and

$$\int_{-\infty}^{\infty} x^2\ f(x)\ dx = F_0\ \sqrt{\pi}\ a_f^{\,3}/2 \tag{6.19}$$

giving

$$a_f^{\,2} = 2 \int_{-\infty}^{\infty} x^2\ f(x)\ dx \Big/ \int_{-\infty}^{\infty} f(x)\ dx. \tag{6.20}$$

Also, if f(x) is Gaussian, the expansion (6.2) (p.106) for F(w) only requires a single term, and as in (6.20) we get

$$a_f^{\,2} = \int_{-\infty}^{\infty} |F(w)|dw \ / \ \int_{-\infty}^{\infty} w^2 |F(w)|dw/2\pi^2 \tag{6.21}$$

Procedure P4 for finding a_f and a_g when M_F and M_G are given.

P4.1 Guess a value for a_f and represent f(x) by the Hermite functions using P3 (p.109)

P4.2 Calculate the root mean squared error from equation (6.17) (p.109).

P4.3 Iteratively adjust a_f in P4.1 until the mean squared error is minimised. The algorithm for iteratively adjusting a_f is given in section 8.3.3.

P4.4 Repeat P4.1-P4.3 to estimate a_g from the data g(x).

P4.5 Calculate $a_h^{\,2} = a_f^{\,2} - a_g^{\,2}$.

Procedure P5 for finding a_f and a_g.

P5.1 Estimate a_f directly from equation (6.20).

P5.2 Estimate a_g directly from equation (6.20) with a_f and f(x) replaced by a_g and g(x) respectively.

P5.3 Calculate $a_h^2 = a_f^2 - a_g^2$.

Procedure P6 for finding a_f and a_g.

P6.1 Calculate the Fourier transform F(w) of f(x) using (5.12) (p.91). (See section 8.1, p.153 for computational details.)

P6.2 Estimate a_f^2 from (6.21).

P6.3 Calculate the Fourier transform G(w) of g(x) as in P6.1.

P6.4 Estimate a_g^2 from (6.21) with a_f^2 and $|F(w)|$ replaced by a_g^2 and $|G(w)|$ respectively.

P6.5 Calculate $a_h^2 = a_f^2 - a_g^2$.

Procedure P7 for finding a_f and a_g when M_F is known.

P7.1 Adjust a_f iteratively to minimise the RMS error as in P4.1, P4.2 and P4.3 (p.111).

P7.2 Compute F(w) and G(w) from f(x) and g(x) as in P6.1 and P6.3.

P7.3 Plot a graph of $|F(w)/G(w)|$.

P7.4 Inspect $|F(w)/G(w)|$, and choose a value $w = w_{max}$ beyond which it diverges due to error in the data. (This is like inverse filtering; cf. section 5.1, p.91).

P7.5 Set $|F(w)/G(w)| = 0$ for $|w| > w_{max}$.

P7.6 Estimate a_h^2 from (6.21) with a_f^2 and $|F(w)|$ replaced by a_h^2 and $|F(w)/G(w)|$ respectively.

P7.7 Calculate $a_g^2 = a_f^2 - a_h^2$.

Procedure P8 for finding a_f and a_g when M_G is known.

P8.1 Adjust a_g iteratively to minimise the RMS error as in P4.1, P4.2 and P4.3 (p.111).

P8.2 Estimate a_h^2 using steps P7.2-P7.6.

P8.3 Calculate $a_f^2 = a_g^2 + a_h^2$.

Of the procedures P4-P8, P5 and P6 have the advantages of being independent of M_F and M_G, and of not requiring iterative computations. However, before we assess (in section 6.3) the various procedures P4-P8 we must consider how to estimate M_F and M_G.

6.1.2 Determining the Number of Zeros

As the number of terms M_F in the representation (6.1) (p.106) of f(x) by Hermite functions is increased, the root mean squared error must decrease. Therefore, if we have an estimate of the experimental error we can find M_F as the number of terms in (6.1) required to obtain a root mean squared error just less than the experimental error.

We obtain a relationship between M_F and M_G if we use the number of available data of f(x) and g(x) as an indication of the amount of 'information' we have about each function. We may constrain M_F/M_G to be equal to the ratio N_f/N_g where N_f and N_g are the number of data of f(x) and g(x) respectively. We note that if the sample interval of the data is the same for both functions, the ratio of their 'widths' will be the same as the ratio N_f/N_g.

The above suggestions are the basis of the first two of the following procedures for determining M_F and M_G when a_f and a_g are known. Procedures P11 and P12 are based directly on the behaviour of the zeros ξ_F and ξ_G as M_F and M_G vary.

Procedure P9 for determining M_F or M_G from an estimate of the experimental error.

P9.1 Estimate the experimental error in the data f(x) (or g(x)).

P9.2 Carry out procedure P3 (p.109) for finding the coefficients of the Hermite function expansion, increasing M_F (or M_G) until the RMS error given by (6.17) (p.109) falls below the estimated experimental error obtained in P9.1.

Procedure P10 for finding a relationship between M_F and M_G.

P10.1 Estimate the intervals L_f and L_g over which the data f(x) and g(x) respectively, have significant values.

P10.2 Constrain the number of zeros so that $M_F/M_G = L_f/L_g$.

Procedure P11 for finding M_F or M_G by observing the convergence of the zeros.

P11.1 Choose a value for M_F (or M_G) and, using P3 (p.109) followed by P1 (p.107), calculate and plot the positions of the zeros ξ_F (or ξ_G).

P11.2 Repeat P11.1 a number of times, increasing M_F (or M_G) by 2 at each stage.

P11.3 Identify the value of M_F (or M_G) beyond which the positions of at least some of the zeros remain essentially constant.

Procedure P12 for finding M_F or M_G by trial.

P12.1 For a fixed value of M_G (or M_F) calculate the zeros ξ_G (or ξ_F) using P3 (p.109) followed by P1 (p.107).

P12.2 Using P3 followed by P1 calculate the zeros ξ_F (or ξ_G) for various $M_F \geq M_G$ (or $M_G \leq M_F$).

P12.3 For each M_F (or M_G) calculate the criterion (5.30) (p.102) which indicates how well the zeros ξ_F and ξ_G correspond.

P12.4 Accept the M_F (or M_G) for which correspondence is best.

We note that procedures P9-P12 are progressively less precise, and that the lack of preciseness must be compensated for by the interaction of the user with the computations. The overall efficiency of the deconvolution procedure will be determined by the particular combination of procedures P4-P8 and P9-P12 which is used. Undoubtedly, deconvolution using any combination of procedures P5 or P6 with P9 or P10, will be more efficient than deconvolution using any of the procedures P4, P7, P8, P11 or P12.

6.2 AN EXAMPLE OF CONSISTENT DATA

To evaluate the various procedures described in sections 6.1.1 and 6.1.2, we attempt the deconvolution of data which is known to be consistent. Fig. 6.1 (p.105) shows f(x) obtained when h(x) (Fig. 6.4) is convolved with g(x) (Fig. 6.2, p.105) by approximating (3.56) by a sum. The particular h(x) in Fig. 6.4 is not known to be exactly represented by a finite number of Hermite functions, and is typical of the h(x) derived from X-ray diffraction experiments (cf. Klug and Alexander, 1967). We firstly discuss those procedures which give a useful deconvolution when applied to the data of Figs 6.1 and 6.2, and subsequently demonstrate the unsatisfactory performance of the remaining procedures.

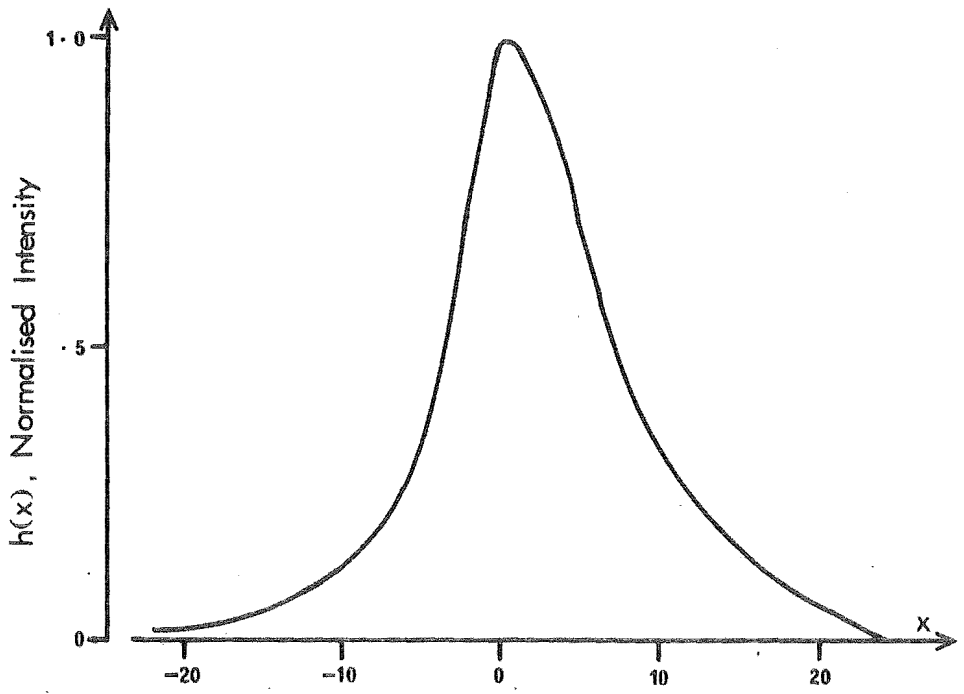


Fig. 6.4 h(x) which was convolved with g(x) in Fig. 6.2 (p.105) to give f(x) in Fig. 6.1.

6.2.1 Results from Successful Procedures

Table 6.5 is a detailed summary of the results obtained when procedures P4, P7 or P8 coupled with P11 or P12 were used to deconvolve the data of Figs 6.1 and 6.2 (p.105). The estimates of a_f, a_g, M_F and M_G are given along with the various Tables and Figures obtained at certain steps in the computations.

We note in Fig. 6.11, which displays the zeros of F(w) and G(w) obtained using P8 (see Table 6.5), that no zero $w_{F,M}$ of F(w) lies in the same position as any zero $W_{G,M}$ of G(w). Our data which we know to be 'consistent' appears 'inconsistent', because of the difficulty in finding suitable values for a_f, a_g, M_F and M_G. We can readily identify (cf. the arrows in Fig. 6.11) members of ξ_F which lie closest to the

Table 6.5

Summary of Results Using Successful Procedures

Estimate a_f			Estimate a_g			Estimate M_F			Estimate M_G			Fig. showing zeros of F(w) & G(w)[1]	Fig. showing h(x) reconstructed from ξ_H[2]	Fig. showing effect of shifting G(w) zeros[3]	Fig. showing effect of shifting F(w) zeros[4]
Procedure	Result	Remarks	Procedure	Result	Remarks	Procedure	Result	Remarks	Procedure	Result	Remarks				
P4 (p. 111)	9.86	RMS error = .3. Difference between data and its Hermite function representation could not be seen on Fig. 6.1.	P4	8.95	RMS error = 3.0. Difference between data and its Hermite function representation could not be seen on Fig. 6.2.	P12	26		P12	14		Fig. 6.8	Fig. 6.12		
P7 (p. 112)	9.86	As for P4.	P7	8.07	Fig. 6.6 shows $\lvert F(w)/G(w) \rvert$ from P7.3. w_{max} from P7.4 is shown as the single-headed arrow. P7.6 gives $a_h = 5.67$.	P12	26		P12	14		Fig. 6.9	Fig. 6.12		
P8 (p. 113)	10.62	Fig. 6.6 shows $\lvert F(w)/G(w) \rvert$. w_{max} is the double-headed arrow. P8.2 gives $a_h = 5.74$.	P8	8.95	As for P4.	P12	24 or 26	P12 did not distinguish between $M_F = 24$ and $M_F = 26$.	P12 or P11	14	Subsequent to using P12, Table 6.7 was calculated using P11.	Fig. 6.10 for M_F = 24. Fig. 6.11 for M_F = 26.	Fig. 6.13 for M_F = 26. Reconstruction for M_F = 24 is very similar.	Fig. 6.14 for M_F = 26.	Fig. 6.15 for M_F = 26.

[1] In the Figures cited, crosses denote zeros of F(w), and circles denote zeros of G(w).

[2] The Figures compare h(x) calculated from the zeros ξ_H (dashed line) with the original h(x) (Fig. 6.4, p.116).

[3] The zeros of G(w) are shifted to coincide with the corresponding F(w) zeros. The Figures compare the resulting $g_f(x)$ (dashed line) with the original g(x) (Fig. 6.2, p.105).

[4] Zeros of F(w) are shifted to coincide with the corresponding G(w) zeros. The Figures compare the resulting $f_g(x)$ (dashed line), which is the convolution of the reconstructed h(x) and the original g(x), with the original f(x) (Fig. 6.1, p.105).

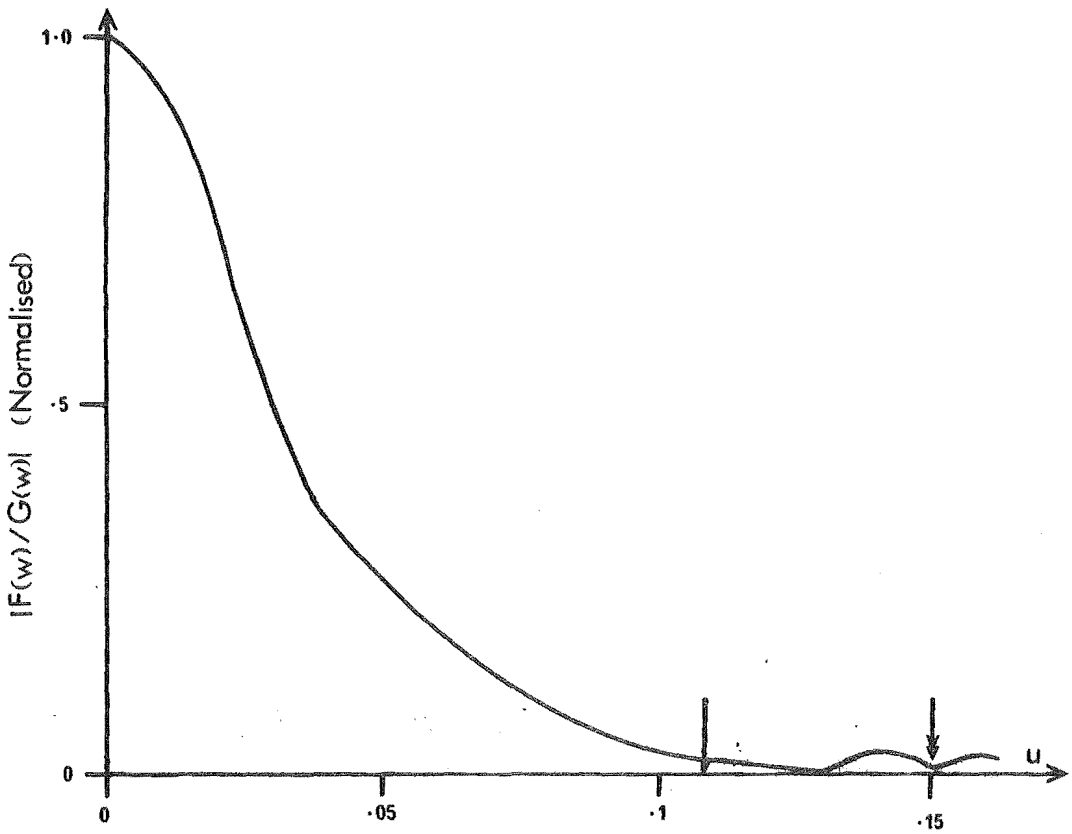


Fig. 6.6 |F(w)/G(w)| computed when applying procedures P7 and P8 to the data in Figs 6.1 and 6.2 (p.105)

Table 6.7

Convergence of the Zeros of G(w) with a_g = 8.95 for Various Values of M_G.

(Only zeros on the imaginary axis and in the right half w-plane are listed)

Positions of Zeros							
M_G = 12		M_G = 14		M_G = 16		M_G = 18	
$U_{G,m}$	$V_{G,m}$	$U_{G,m}$	$V_{G,m}$	$U_{G,m}$	$V_{G,m}$	$U_{G,m}$	$V_{G,m}$
0.	-0.1369						
0.	0.1034						
0.0241	-0.0282	0.0194	0.0234	0.0195	0.0235	0.0193	0.0235
0.0343	-0.0071	0.0347	-0.0072	0.0347	-0.0072	0.0347	-0.0071
0.0421	-0.0345	0.0323	-0.0393	0.0321	-0.0307	0.0295	-0.0299
0.0667	0.0277	0.0548	0.0246	0.0549	0.0247	0.0548	0.0252
		0.0808	-0.0369	0.0802	-0.0363	0.0723	-0.0361
		0.0917	0.0261	0.0919	0.0264	0.0931	0.0278
0.1010	-0.0082	0.1032	-0.0063	0.1032	-0.0062	0.1028	-0.0065
				0.9799	0.2046		
						0.1088	0.0846
						0.1327	-0.0447

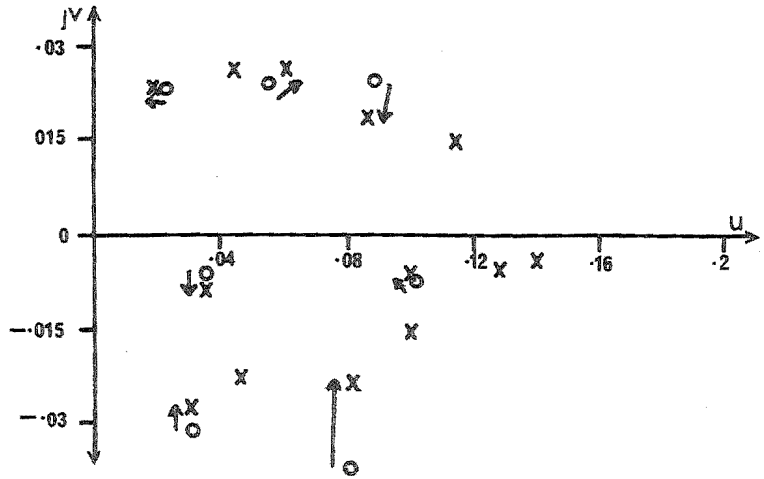


Fig. 6.8 Zeros of F(w) (XXX) and zeros of G(w) (OOO) computed using procedure P4.

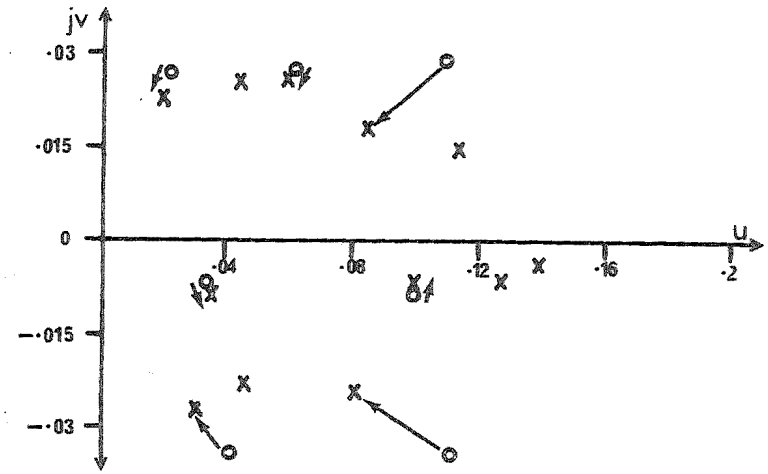


Fig. 6.9 Zeros of F(w) (XXX) and zeros of G(w) (OOO) computed using procedure P7

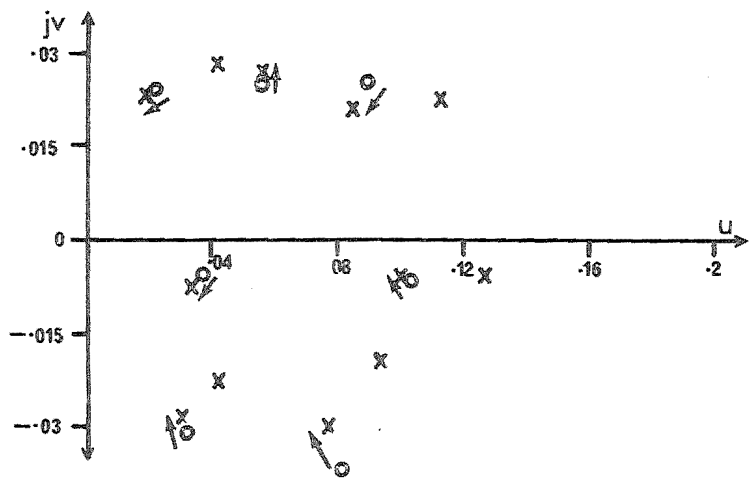


Fig. 6.10 Zeros of F(w) (XXX) and zeros of G(w) (OOO) computed using procedure P8 with $M_F = 24$.

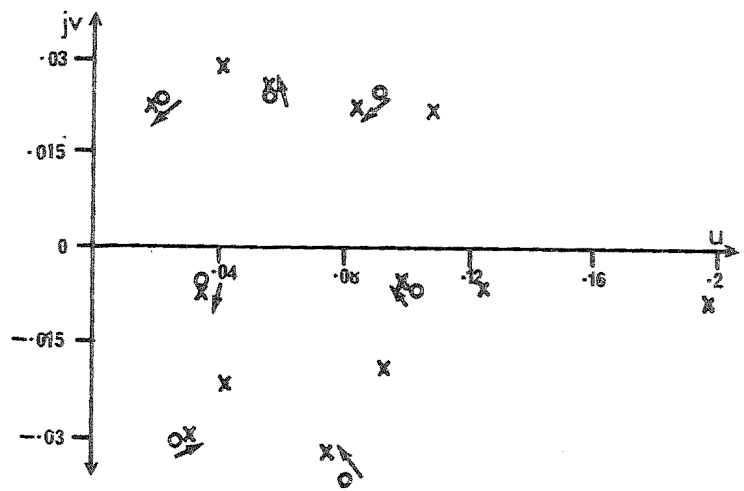


Fig. 6.11 Zeros of F(w) (XXX) and zeros of G(w) (OOO) computed using procedure P8 with $M_F = 26$.

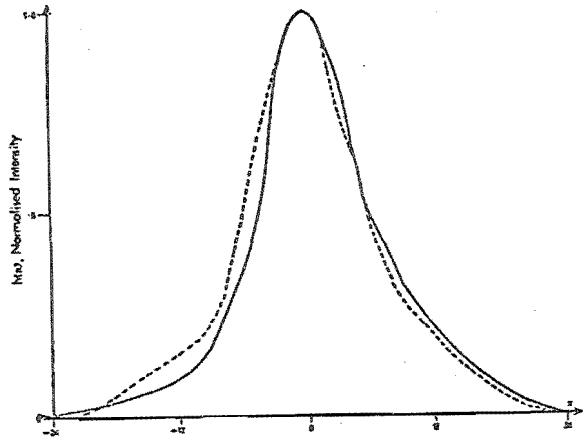


Fig. 6.12 h(x) reconstructed from the unmatched zeros of Fig. 6.8 or Fig. 6.9 (----), compared with the original h(x) (———).

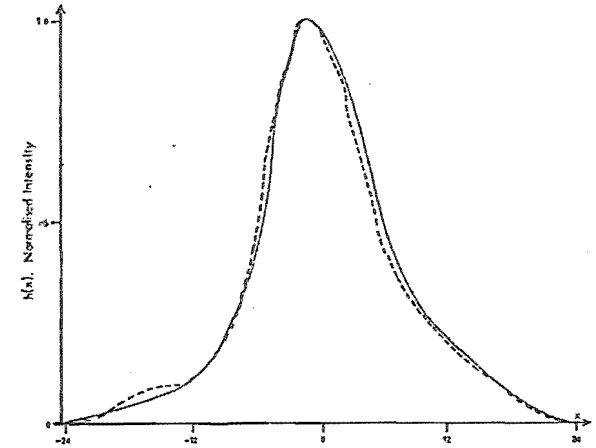


Fig. 6.13 h(x) reconstructed from the unmatched zeros of Fig. 6.11 (----), compared with original h(x) (———).

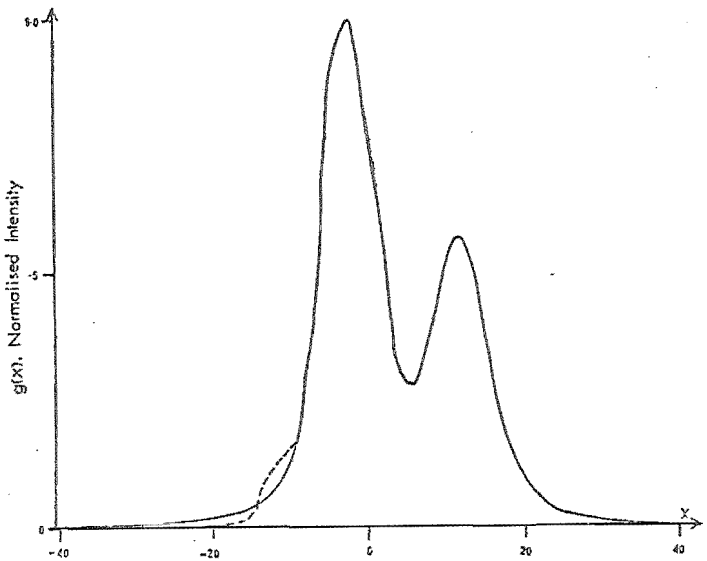


Fig. 6.14 $g_g(x)$ obtained by shifting the zeros of G(w) in Fig. 6.11 to coincide with those of F(w) (----), compared with the original g(x) (———).

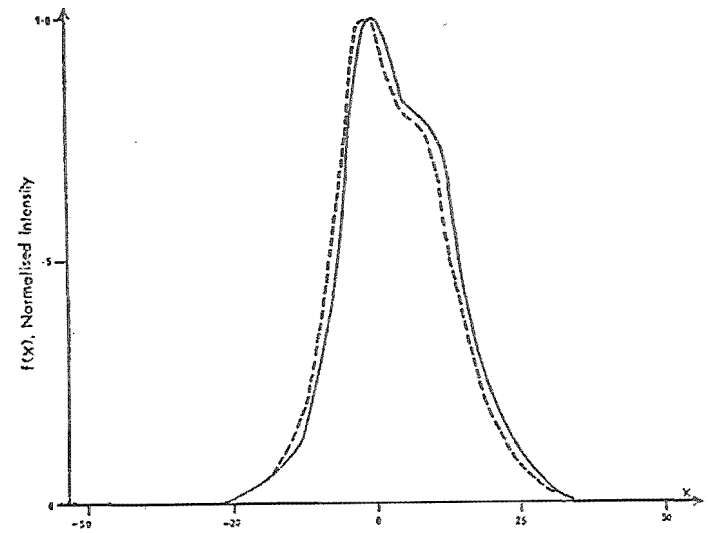


Fig. 6.15 $f_g(x)$ obtained by shifting the zeros of F(w) in Fig. 6.11 to coincide with those of G(w) (----), compared with the original f(x) (———)

members of ξ_G in the sense defined by equation (5.30) (p.102). Therefore the crosses in Fig. 6.11 which are not marked by an arrow, are the zeros ξ_H of H(w). We have $M_H = 12$ because there are six further unmatched ξ_F zeros in the left half w-plane not shown in Fig. 6.11. Using procedure P2 (p.107), the deconvolved h(x) shown as the dashed line in Fig. 6.13 is computed from ξ_H.

To obtain the h(x) in Fig. 6.13, we have implicitly assumed that the zeros of F(w) are more accurate than those of G(w). Therefore, we generate a new 'consistent' G(w) by shifting its zeros in the direction of the arrows in Fig. 6.11. This procedure is equivalent to the zero shifting operation described by equation (5.32) (p.102), and gives the set of zeros ξ_{GF} of the function $G_F(w)$. The corresponding $g_f(x)$ obtained using P2 (p.107) on the zeros ξ_{GF}, is shown in Fig. 6.14 together with the original g(x) reproduced from Fig. 6.2.

If we shift the zeros of ξ_F in the opposite direction to the arrows in Fig. 6.11, we carry out the zero shifting procedure described by equation (5.33) (p.103) and obtain the zeros ξ_{FG} of $F_G(w)$. Because

$$\xi_{FG} = \xi_G \cup \xi_H, \tag{6.23}$$

the reconstruction of $f_g(x)$ from the zeros ξ_{FG} using P2 (p.107) is equivalent to the convolution of the reconstructed h(x) shown in Fig. 6.13 (dashed line) with the original g(x) (Fig. 6.2, p.105). We compare $f_g(x)$ with f(x) in Fig. 6.15. In section 5.2.2 (p.103) we suggested that it is unreasonable to assume that only the first M_G zeros in ξ_F are in error, and that a more 'consistent' f(x) would be obtained by shifting the matched members of ξ_F as above, and then adjusting the

unmatched members to give some sort of best fit to f(x). We have not attempted this procedure, which is a nonlinear optimisation problem.

Although the reconstructed h(x) in Fig. 6.13 is in good agreement with the original, the procedures used to obtain it are, computationally, among the least efficient of those in sections 6.1.1 and 6.1.2. A large amount of human interaction with the computations is required, which means that there is always a possibility of not finding suitable values for a_f, a_g, M_F and M_G. Procedure P4, requiring optimisation of both a_f and a_g is less efficient than P7 and P8 which require optimisation of only a_f or a_g. In section 6.3 we attempt the deconvolution of experimental data using procedures P8 and P12, but firstly we demonstrate the unsatisfactory values for a_f, a_g, M_F and M_G obtained using procedures P5 and P6 (p.112) and P9 and P10 (p.114).

6.2.2 Results from Unsuccessful Procedures

Table 6.16 is a summary of results obtained using the procedures in sections 6.1.1 and 6.1.2 which were abandoned for various reasons. The information given in Table 6.16 does not prove that the procedures listed are always bound to fail. Rather, it shows the inferior results obtained during tests on one particular set of data.

Table 6.16

Summary of Results from Unsuccessful Procedures for Finding a_f, a_g, M_F and M_G Compared With the 'Successful' Values of $a_f = 10.62$, $a_g = 8.95$, $M_F = 26$, $M_G = 14$.

Procedure	Results	Remarks
P5(p.112)	P5.2 gives $a_g = 13.6$	Table 6.17 shows the relatively poor convergence of the Hermite representation of g(x) when $a_g = 13.6$ compared to the convergence when $a_g = 8.95$.
P6(p.112)	P6.2 gives $a_f = 11.3$ P6.4 gives $a_g = 5.66$	Table 6.17 shows the relatively poor convergence of the Hermite representation of g(x) when $a_g = 5.66$ compared to the representation when $a_g = 8.95$.
P9(p.114)	Various	We note from Table 6.17 that with $a_g = 8.95$ the root mean squared error changes from 4.7 to 2.3 as M_G increases from 12 to 20. It is unlikely that an estimate of experimental error would be sufficiently accurate to distinguish among a number of possible values for M_G.
P10(p.114)	P10.2 gives $M_F/M_G = 107/80 = 1.35$	Choosing the previously successful $a_g = 8.95$, $M_G = 14$, and $a_f = 10.62$ we get $M_F = 18$. The resulting zeros plotted in Fig. 6.18a do not correspond as well as those in Fig. 6.18b which is a reproduction of Fig. 6.11. Any attempt to reconstruct h(x) from unmatched zeros in Fig. 6.18a resulted in a function with large negative values.

Table 6.17

Convergence of the Hermite Function Representation of g(x) (Fig. 6.2, p.105) for Various Values of a_g.

Order of Hermite function representation (m)	RMS error after m terms		
	a_g = 13.6 (from P5)	a_g = 8.95 (from P8)	a_g = 5.66 (from P6)
0	101.	111.	139.
1	99.3	110.	126.
2	99.1	104.	108.
3	88.5	45.4	95.6
4	83.2	23.8	80.4
5	71.0	12.7	64.8
6	60.2	11.4	45.0
7	49.8	10.1	35.1
8	42.0	9.9	23.9
9	32.1	9.9	18.5
10	25.9	5.0	14.4
11	20.0	4.9	13.2
12	14.7	4.7	10.8
13	12.4	4.6	10.1
14	9.2	2.5	8.9
15	8.7	2.5	8.6
16	7.3	2.5	7.4
17	7.3	2.5	7.2
18	7.0	2.4	6.5
19	6.7	2.3	6.5
20	6.6	2.3	5.8

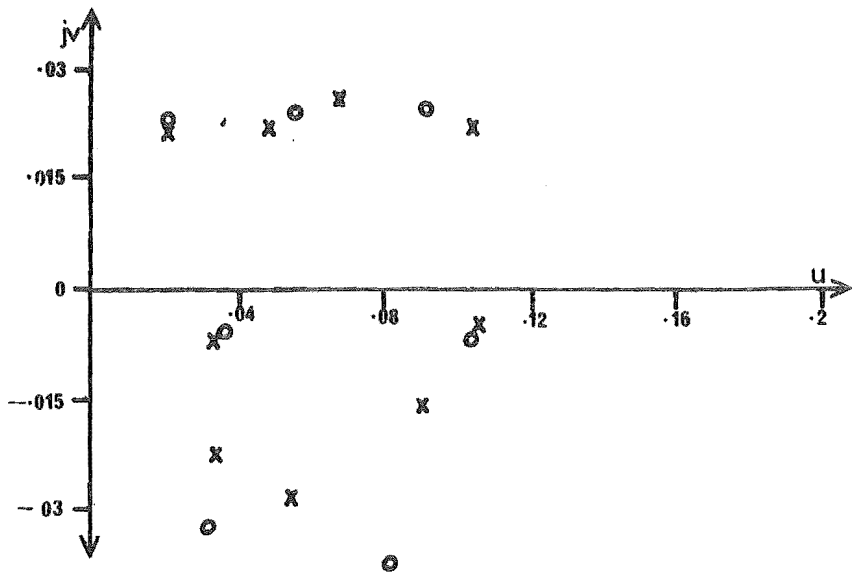


Fig. 6.18a Zeros of F(w) (XXX) and zeros of G(w) (OOO) computed using $a_f = 10.62$, $a_g = 8.95$, $M_F = 18$ and $M_G = 14$.

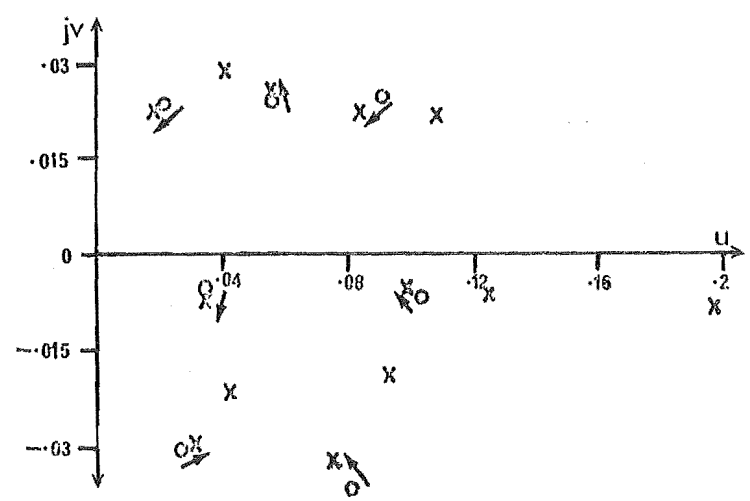


Fig. 6.18b Zeros of F(w) (XXX) and zeros of G(w) (OOO) computed using $a_f = 10.62$, $a_g = 8.95$, $M_F = 26$, and $M_G = 14$. (Reproduced from Fig. 6.11.)

6.3 DECONVOLUTION OF X-RAY DIFFRACTION DATA

X-ray diffraction measurements have proved useful in the study of various properties of solids (cf. Klug and Alexander, 1967). However, typical apparatus distorts (broadens) the diffraction pattern, making quantitative deductions difficult. For small angle scattering, the desired pattern h(x) is distorted by the 'instrument function' g(x) to give the measured pattern f(x). h(x), g(x) and f(x) are related by the convolution integral (3.56) (p.61) (cf. Klug and Alexander, 1967, p.494). Our attention was drawn to this problem by Kivell (1973), who was attempting to predict the fatigue life of aluminium samples, from X-ray diffraction measurements. Although we also tried our methods on data from Kivell (1973), we report here our attempts to deconvolve the similar data published by Stokes (1948).

Figures 6.19 and 6.20 show the measured f(x) and g(x) respectively, and in Table 6.21 we indicate the various results obtained during the attempted deconvolution.

The zeros obtained from the results in Table 6.21 are shown in Fig. 6.23, where we note that the correspondence between members of ξ_F and ξ_G is poor compared with our 'consistent' data (Fig. 6.11, p.120). Procedure P11 (p.114) does not increase our confidence in the values of a_f, a_g, M_F and M_G given in Table 6.21, as neither the zeros of G(w) nor F(w) showed any sign of convergence with varying M_G and M_F respectively.

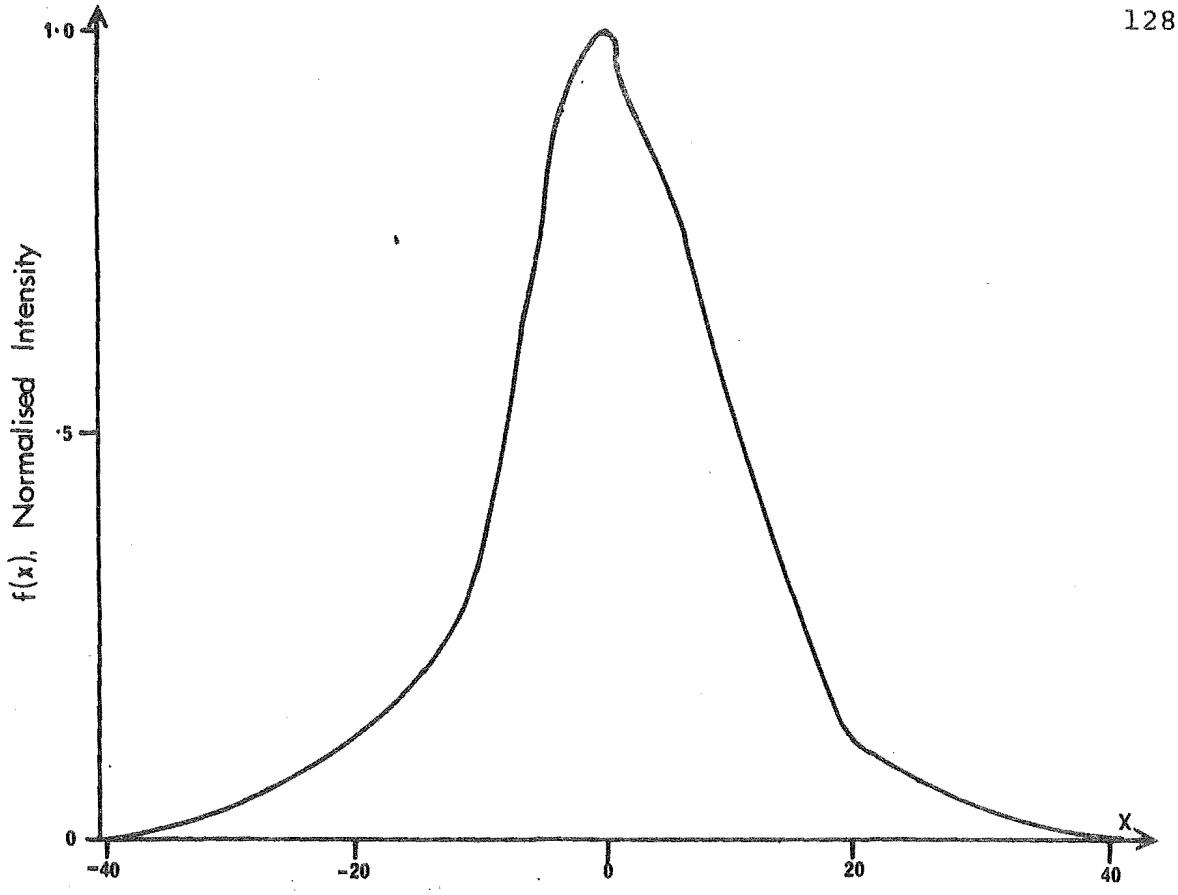


Fig. 6.19 Measured f(x) from Stokes (1948).

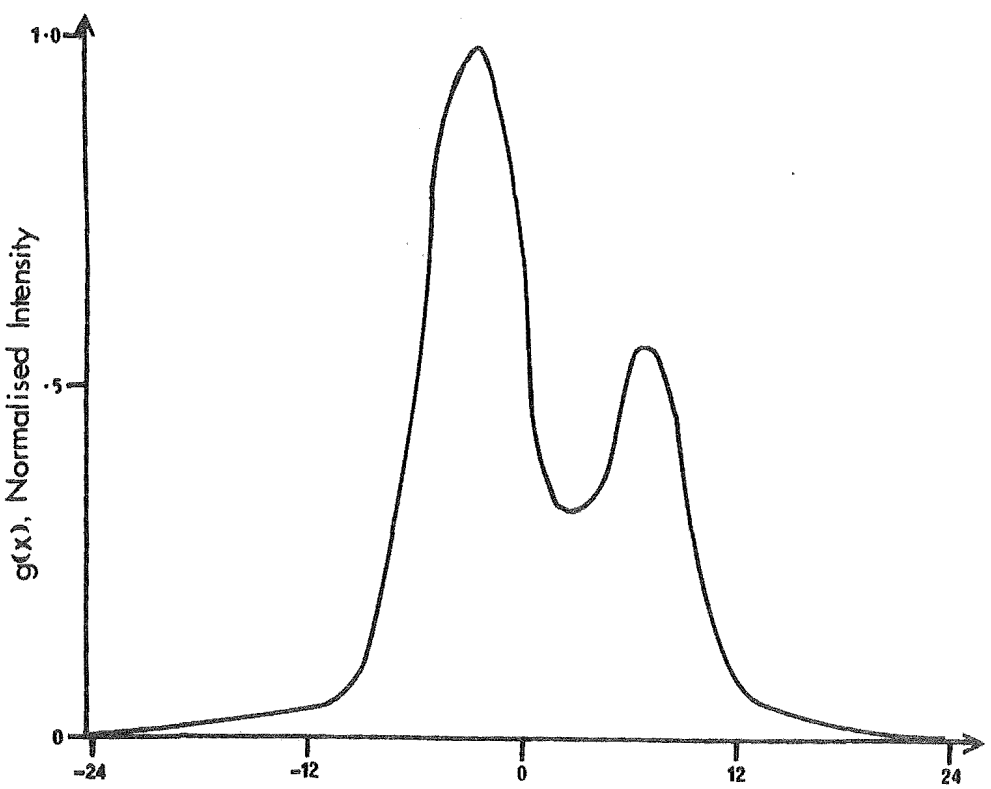


Fig. 6.20 Measured g(x) from Stokes (1948).

Table 6.21

Summary of Steps in the Deconvolution of Data from Stokes (1948)

Estimate a_f	Estimate a_g	Estimage M_F	Estimate M_G
Fig. 6.22 shows $\|F(w)/G(w)\|$ from P8.2. w_{max} is shown by the arrow. P8.2 gives $a_h = 5.5$ so that from P8.3, $a_f = 6.94$.	Using P8 we get $a_g = 4.27$ corresponding to an RMS error of 2.5.	P12 gives $M_F = 28$.	P12 gives $M_G = 20$.

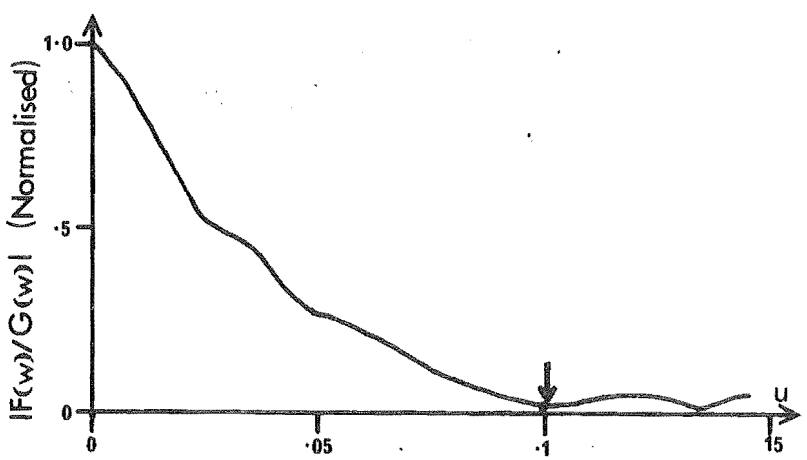


Fig. 6.22 $|F(w)/G(w)|$ obtained when applying procedure P8.2 to the data of Figs. 6.19 and 6.20.

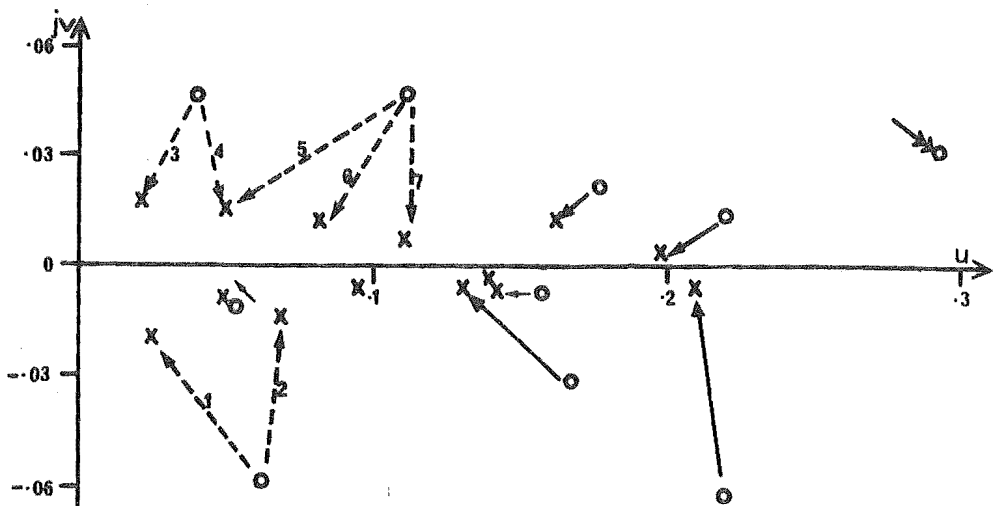


Fig. 6.23 Zeros of F(w) (XXX) and zeros of G(w) (OOO) computed from the data of Fig. 6.19 and 6.20.

Notwithstanding, the arrows in Fig. 6.23 indicate attempts to match the zeros of G(w) with those of F(w). Where there is ambiguity, dashed, numbered arrows indicate the various possibilities. We cannot match the zero of G(w) marked with a double-headed arrow in Fig. 6.23. Although the presence of this zero and its counterpart in the left half w-plane has a noticeable effect on g(x) (Fig. 6.24), we omit the offending pair from ξ_G and attempt to identify the set ξ_H from the remaining zeros.

Because of the ambiguity evident in Fig. 6.23, there are various possible sets ξ_H. We require h(x) to be both positive and approximating a Gaussian function (cf. Klug and Alexander, 1967). The h(x) best satisfying these requirements is shown in Fig. 6.25 (dashed line) which is reconstructed from ξ_H obtained when the zeros in Fig. 6.23 are matched according to the solid arrows and the dashed arrows numbered 1, 3 and 5.

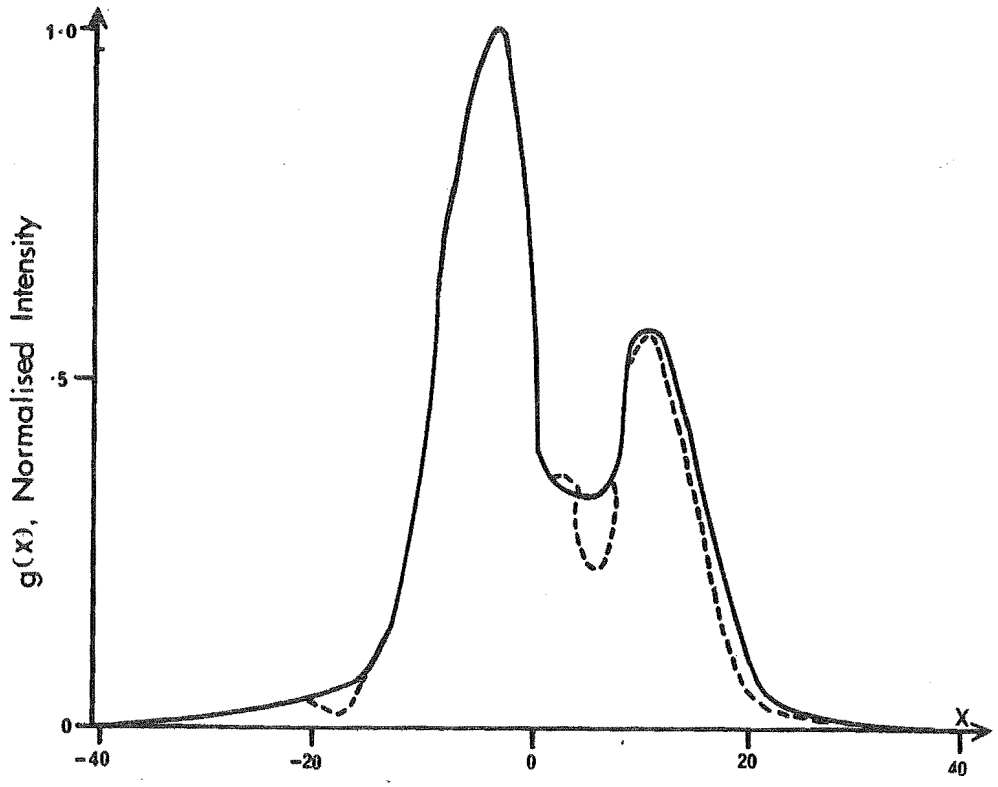


Fig. 6.24 The effect of g(x) on the zero of G(w) marked with a double-headed arrow in Fig. 6.23. ---- g(x) reconstructed without the offending zeros; —— original g(x) (Fig. 6.20, p.128).

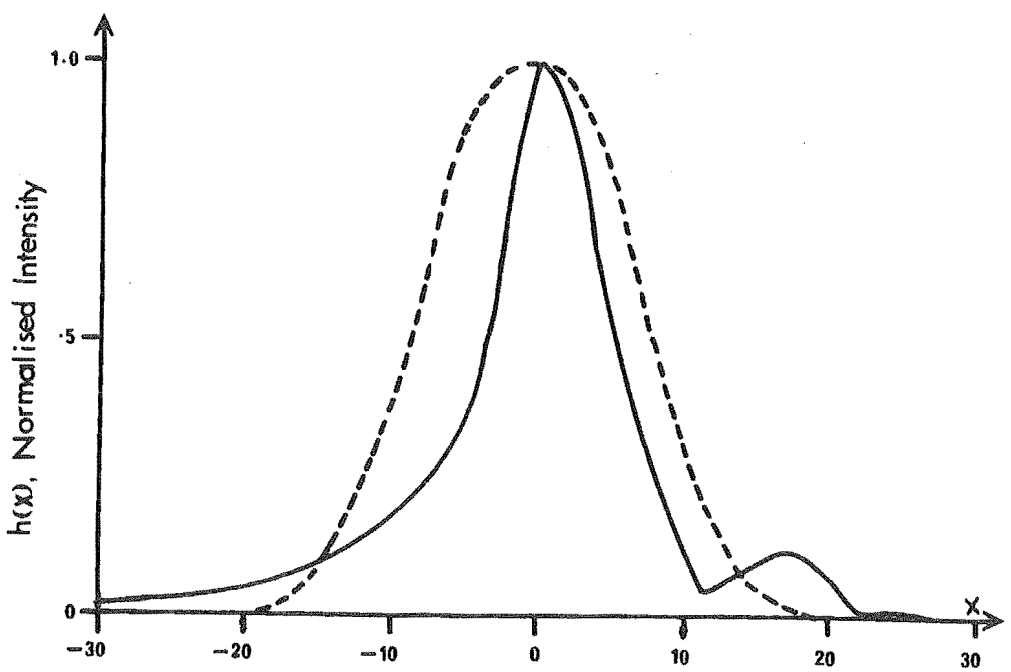


Fig. 6.25 ---- denotes the best h(x) obtained by matching the zeros in Fig. 6.23. —— denotes h(x) obtained by inverse filtering.

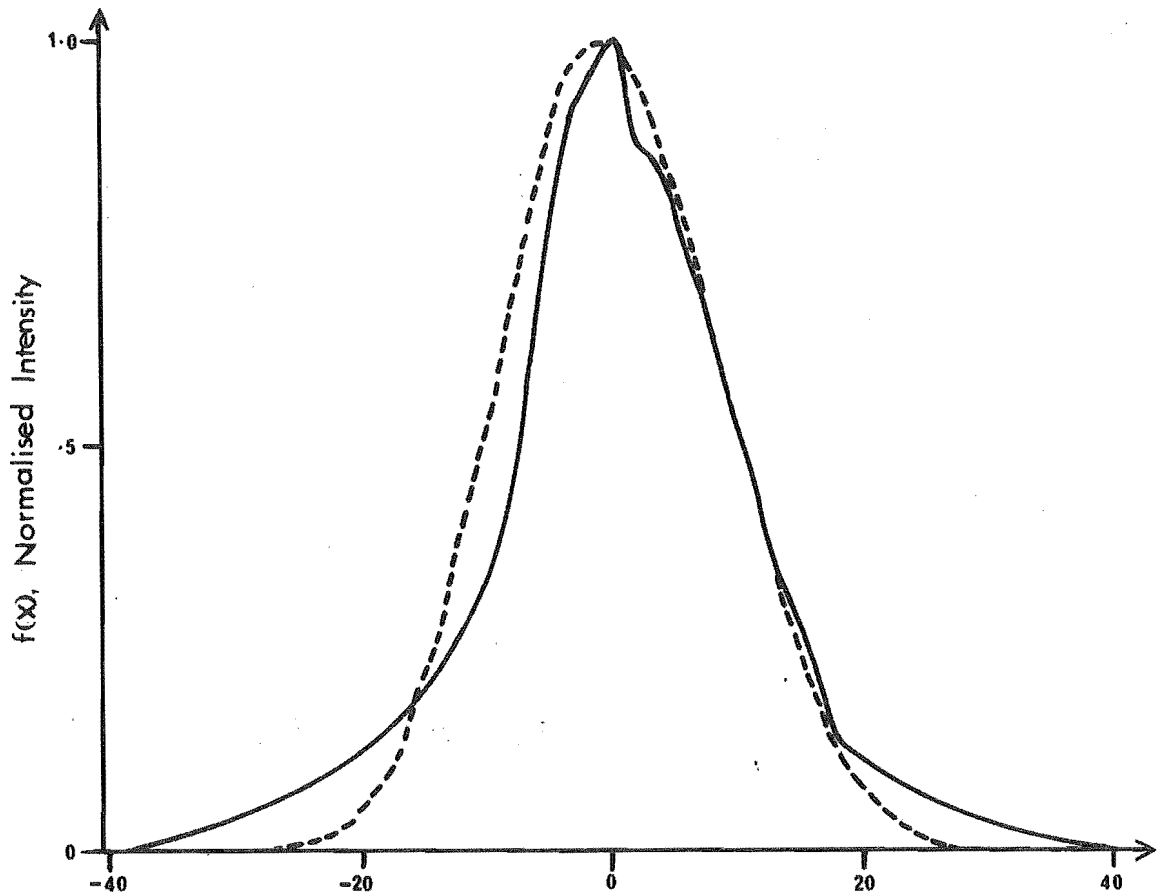


Fig. 6.26 ---- denotes h(x) (Fig. 6.25 dashed line) convolved with the original g(x) (Fig. 6.20, p.128).
—— denotes the original f(x) (Fig. 6.19)

In Fig. 6.26 we see that the convolution of the reconstructed h(x) with the original g(x) is much wider than the original f(x); further evidence to suggest that we can have little confidence in the reconstructed h(x) of Fig. 6.25. The data f(x) (Fig. 6.19, p.128) and g(x) (Fig. 6.20) have previously been deconvolved by inverse filtering (Stokes, 1948) to obtain the h(x) shown as the continuous line in Fig. 6.25. The latter result is regarded as useful (cf. Stokes, 1948), although the irregularity on the right side of the main peak is not given any physical interpretation.

Although we have reported attempts to deconvolve the data of Stokes (1948) using only procedures P8 and P12, no more success was achieved with various of the other procedures described in sections 6.1.1 and 6.1.2. We conclude that our approach to the deconvolution of complete data has been unsuccessful. The difficulties in estimating suitable values of a_f, a_g, M_F and M_G have not been overcome, so that the potential of the condition (5.28) (p.101) to discover and correct, inconsistent data has not been realised.

C H A P T E R 7

DECONVOLUTION OF INCOMPLETE DATA

The deconvolution of data known or estimated with sufficient accuracy throughout the range $-\infty$ to $+\infty$ is described in chapter 6. The techniques of this chapter deal with 'incomplete' data, that is data which is available in only a stated range, and which is completely unknown outside that range. Methods proposed for solving this Type B problem are reviewed in section 3.4.1. Fourier methods are not usually applied directly to the deconvolution of truncated data, because the effect of the truncation on the accuracy of any Fourier transform is more serious than is the effect of the experimental error.

A typical deconvolution problem with incomplete data is shown in Fig. 7.1 where we have f(x) (continuous line) given by

$$f(x) = e^{-0.5x}\left[\frac{e^{-0.3x} - 1}{0.3} + \frac{e^{-0.4x} - 1}{0.4}\right] \qquad (x > 0). \quad (7.1)$$

Equation (7.1) is the convolution of

$$g(x) = e^{-0.5x} \qquad (x > 0) \qquad (7.2)$$

with

$$h(x) = e^{-0.2x} - e^{-0.9x} \qquad (x > 0). \qquad (7.3)$$

In this case f(x) and g(x) exist for all positive x, but we take them as known only in the interval $0 \leqslant x \leqslant 6.0$. Inspection of the convolution integral (3.56) (p.61) shows that if the values of f(x) in the interval $0 \leqslant x \leqslant L_g$ are

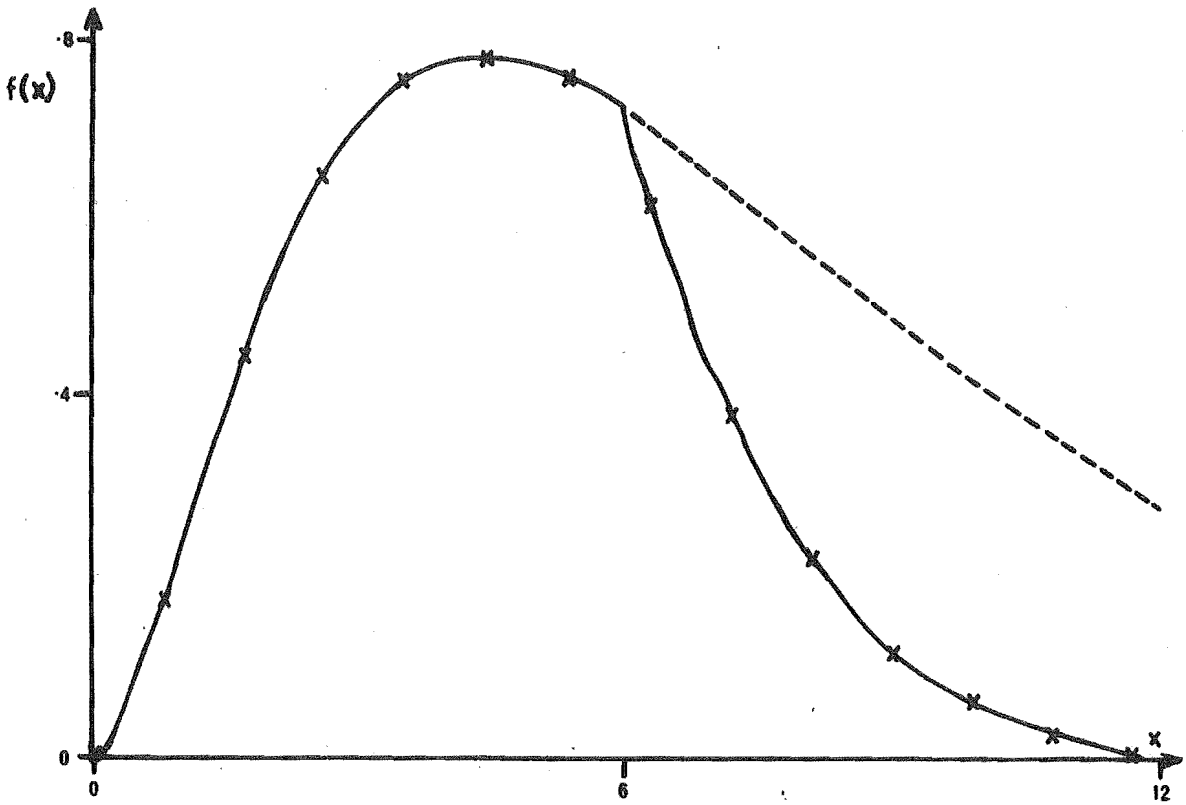


Fig. 7.1 f(x), the convolution of g(x) (7.2) and h(x) (7.3). ×——× g(x) and h(x) of extent L_g = 6; ---- g(x) and h(x) of infinite extent.

known, and the function g(x) is known in the same interval, the function h(x) is also completely determined in the interval, but not outside it. Therefore, for functions f(x) and g(x) given in the interval $0 \leqslant x \leqslant L_g$ without error, we can approximate the convolution integral (3.56) by a sum, and find h(x) by solving the equations (5.8) (p.90). However, measured functions f(x) and g(x) contain error, the effects of which we wish to reduce using the filtering properties of the Fourier transform (cf. section 5.1.).

Because the values of g(x) and h(x) outside the interval $0 \leqslant x \leqslant L_g$ are unknown, and do not affect f(x) inside the interval, we choose g(x) and h(x) to be zero unless $0 \leqslant x \leqslant L_g$.

Applying this constraint to g(x) and h(x) their convolution f(x), for our example (7.1), (7.2) and (7.3), is shown as the line with crosses in Fig. 7.1. Now f(x) is 'complete'. We note that in the range $6 < x \leqslant 12$ the completed f(x) does not correspond with the true f(x) (dashed line), but in practice we do not know the behaviour of f(x) in the interval $6 < x \leqslant 12$ anyway. Fourier methods can now be applied to deconvolve the completed f(x) and the g(x) of finite extent.

In a practical problem h(x) is unknown, and so f(x) cannot be completed as described above. Present Fourier methods, effectively complete f(x) by iteratively correcting an initial guess to it in the adjacent interval $L_g \leqslant x \leqslant 2L_g$ (cf. Silverman and Pearson, 1973). We attempt to complete f(x) by using the consistency condition (5.28) (p.101)

7.1 CALCULATION OF ZEROS USING FOURIER SERIES

We have defined g(x), h(x) and the completed f(x) to be of extent L_g, L_h $(= L_g)$, and L_f $(= 2L_g)$ respectively. Therefore, f(x), g(x) and h(x) satisfy the condition (5.18) (p.98) so that their Fourier transforms F(w), G(w) and H(w) are entire functions of order 1 (cf. section 5.2, p.99).

A zero of a function is defined as a value of its argument for which the function has the value zero. A zero of F(w), $w_{F,m} = U_{F,m} + jV_{F,m}$, identifies a point (or position) with coordinates $(U_{F,m}, V_{F,m})$ in the complex w-plane at which $F(w) = 0$.

To calculate the positions of the zeros of the Fourier transform of a function f(x) which is of finite extent $0 \leqslant x \leqslant L_f$, we follow Bates (1969b). A function f(x) which

exists in a finite interval can be represented to any desired accuracy by the Fourier series

$$f(x) = \sum_{m=-m_f}^{m_f} F_m \, e^{-j2\pi mx/L_f} \qquad 0 \leqslant x \leqslant L_f \tag{7.4}$$

where the F_m are complex constants and m_f is a positive integer such that the accuracy of the representation of $f(x)$ by (7.4) increases as m_f increases. Substituting (7.4) into (5.12) (p.91) gives

$$F(w) = e^{j\pi wL_f} \frac{\sin(\pi wL_f)}{L_f \pi} \sum_{m=-m_f}^{m_f} \frac{F(\frac{m}{L_f})}{(w - \frac{m}{L_f})} , \tag{7.5}$$

in which

$$F(\frac{m}{L_f}) = L_f \, F_m . \tag{7.6}$$

we define

$$Q_F(w) = \frac{e^{j2\pi wL_f} \sin(\pi wL_f)}{\pi L_f \prod_{m=-m_f}^{m_f} (w - \frac{m}{L_f})} \tag{7.7}$$

and

$$P_F(w) = \sum_{n=-m_f}^{m_f} F(\frac{n}{L_f}) \prod_{m=-m_f}^{m_f(n)} (w - \frac{m}{L_f}) \tag{7.8}$$

so that

$$F(w) = Q_F(w) \, P_F(w) . \tag{7.9}$$

In (7.8) $\prod_{m}^{(n)}$ denotes a product over m with the term m=n omitted. $Q_F(w)$ is zero in the finite part of the w-plane for only real values of w, for which $u = n/L_f$, where n is an

an integer such that $|n| > m_f$. There are an infinite number of zeros of $Q_F(w)$. $P_F(w)$ is a polynomial of order $2m_f (= M_F)$ which has zeros $w_{F,m}$ such that

$$P_F(w) = C_F \prod_{m=1}^{M_F} (w - w_{F,m}), \tag{7.10}$$

where

$$C_F = (-1)^{m_f} F(0) \prod_{m=1}^{m_f} \left(\frac{m}{L_f}\right)^2 / \prod_{m=1}^{M_F} (w_{F,m}). \tag{7.11}$$

$P_F(w)$ has only a finite number (M_F) of zeros which occur usually for complex values of w (we call them 'complex zeros'), as opposed to the zeros of $Q_F(w)$ which occur only for real values of w (we call these 'real zeros'). The M_F zeros of $P_F(w)$ therefore comprise the set ξ_F, whereas the set of zeros ζ_F consists of the zeros of $P_F(w)$ and the zeros of $Q_F(w)$, and therefore this set has an infinite number of members.

In equation (7.5), we note that F(w) is completely determined by its values $F(\frac{n}{L_f})$ at the sample points $\frac{n}{L_f}$. This is the well known Fourier sampling theorem. Also, (7.8) indicates that $P_F(w)$ depends on the value of F(w) only at its sample points, so that the zeros $w_{F,m}$ in (7.10) are simply an alternative representation of the information in the samples of F(w).

We refer to the following procedure for calculating the zeros ξ_F of the Fourier transform of f(x).

Procedure P13: to calculate the zeros of F(w).

P13.1 Calculate the $2m_f+1$ Fourier coefficients F_m in equation (7.4) using the techniques described in section 8.4.1.

P13.2 Calculate the $2m_f+1$ complex coefficients of the polynomial $P_F(w)$ using (7.8). (A more tractable form of (7.8) is given in section 8.4.2.)

P13.3 Calculate the M_F zeros of $P_F(w)$ as described in section 8.5.

Given its extent L_f, the zeros ξ_F, and the constant C_F, f(x) is reconstructed from the zeros of F(w) as follows:

Procedure P14: to reconstruct f(x) from the zeros of F(w).

P14.1 Compute F(w) using (7.10), (7.7) and (7.9) (p.136).

P14.2 Calculate f(x) using the inverse Fourier transform (5.13) (p.91) as indicated in section 8.1 (p.155).

As in chapter 6, P13 and P14 apply also for the calculation of and reconstruction from the zeros ξ_G and ξ_H, making the appropriate changes in notation.

If f(x) and g(x) are represented exactly by a finite number of terms in a Fourier series of the form (7.4) (p.136), then f(x) requires, in principle, a Fourier series with an infinite number of terms to represent it. We now show this by demonstrating that if m_f is given a finite upper bound, there must be circumstances where $F(w) = 0$, and at the same time $G(w)H(w) \neq 0$. Now since $F(w) = G(w)H(w)$ (cf. equation (5.14), p.91) this is clearly absurd. The particular circumstances arise as follows.

If m_g and m_h are finite, then the coefficients G_m and H_m in the Fourier series representations of g(x) and h(x) are zero for $|m| > m_g$ and $|m| > m_h$, respectively. Therefore, making appropriate changes in notation in (7.6), $G(\frac{2m}{2L_g}) = 0$ for $|m| > m_g$, and $H(\frac{2m}{2L_g}) = 0$ for $|m| > m_h$, and so the product $G(w)H(w)$ is zero for $|m| > m_s$ in $w = \frac{2m}{2L_g}$, where m_s is the smaller of m_g and m_h. Similarly, from the Fourier series

representation of $f(x)$, we have $F(\frac{m}{2L_g}) = 0$ for $|m| > m_f$. Now $F(w) = 0$ at $w = \frac{m}{2L_g}$ for $|m| > m_f$ and $G(w)H(w) = 0$ at $w = \frac{2m}{2L_g}$ for $|m| > m_s$, so that there are real values of w (namely $w = \frac{2m+1}{2L_g}$ for $|2m+1| \geq m_f$) for which $F(w) = 0$, but $G(w)H(w) \neq 0$. This is impossible. If m_f is infinite, there are no values of w for which $F(w) = 0$ but $G(w)H(w) \neq 0$.

In practice, the number of Fourier coefficients which can be calculated from data $f(x)$ is determined by the well-known Fourier sampling criterion

$$m_f = L_f/2T \tag{7.12}$$

where T is the interval in x between successive data points.

The zeros ξ_F calculated from samples of a function $f(x)$ when (7.12) applies are the best available estimates of the zeros of $F(w)$. The accuracy of the estimates improves as m_f increases.

7.2 CONSISTENT COMPLETION AND DECONVOLUTION OF DATA

We aim to complete $f(x)$ so that its Fourier transform $F(w)$ and the Fourier transform of $g(x)$ are such that the consistency condition (5.28) (p.101) holds. Equation (5.28) requires that the zeros ζ_G of $G(w)$ must also be members of ζ_F, the set of zeros of $F(w)$. Therefore, as well as having measurements of $f(x)$ for $0 \leq x \leq L_g$ we know that

$$F(w_{G,n}) = 0 \qquad \text{for all } w_{G,n} \text{ in } \xi_G, \tag{7.13}$$

When $f(x)$ is represented as a Fourier series (7.4) (p.136), equation (7.13) implies that (cf. equation (7.5))

$$e^{j2\pi w_{G,n} L_f} \frac{\sin(\pi w_{G,n} L_f)}{\pi} \sum_{m=-m_f}^{m_f} \frac{F_m}{(w_{G,n} - \frac{m}{L_f})} = 0$$

$$\text{for all } w_{G,n} \text{ in } \xi_G . \qquad (7.14)$$

We obtain a further set of equations from the known values $f(x_n)$ in the range $0 \leq x_n \leq L_g$:

$$\sum_{m=-m_f}^{m_f} F_m e^{-j2\pi m x_n / L_f} = f(x_n) . \qquad (7.15)$$

Equation (7.14) and (7.15) form the basis of

Procedure P15: for completing f(x).

P15.1 Inspect the measured f(x) to find the extent $0 \leq x \leq L_g$ for which it is given, such that $(N_f+1)T = L_g$ where N_f+1 is the number of data points with sample interval T.

P15.2 Assuming the same sample interval for the data g(x), find the $N_f+1 = 2m_g+1$ coefficients in its Fourier series representation similar to (7.4) (p.136), but with f(x), F_m and L_f replaced by g(x), G_m and L_g respectively. Details for finding the Fourier series coefficients are given in section 8.4.1.

P15.3 Find the $M_G = 2m_g$ zeros ξ_G of G(w) using P13 (p.137).

P15.4 Combine the M_G equations (7.14) and N_f+1 equations (7.15) and solve for the $2m_f+1$ coefficients F_m, where $2m_f+1 = M_G + N_f+1$. Details of the form in which (7.14) and (7.15) are expressed and solved are given in sections 8.4.1 (p.165) and 8.6.

P15.5 Plot the completed f(x) after substituting the derived coefficients F_m into equation (7.4) (p.136).

If we now calculate the zeros ξ_F using P13 (p.137), M_G of them are bound by equation (7.14) to coincide exactly with the members of ξ_G. The remaining members of ξ_F comprise ξ_H, from which h(x) can be reconstructed using P14 (p.138). However, we have found it computationally more efficient to deconvolve the completed f(x) and truncated g(x) by inverse filtering (cf. section 3.4.1, p.64, and section 5.1, p.91).

Procedure P16: to deconvolve the completed data.

P16.1 Using equation (5.12) (p.91) with appropriate changes in notation, calculate the Fourier transforms of g(x) and the completed f(x).

P16.2 Plot $|F(u)/G(u)|$, and where it begins to diverge due to error in f(x) or g(x), apply a filter to reduce it to zero (cf. section 8.1, p.155).

P16.3 Calculate h(x) from equation (5.13) (p.91).

7.3 TESTS WITH COMPUTER-GENERATED DATA

We use the example (7.1), (7.2) and (7.3) (p.133) to assess procedures P15 and P16 for completing f(x) and then deconvolving to find h(x). We assume that a measurement of f(x) is available only for $0 \leq x \leq 6.0$ as indicated in Fig. 7.1 (p.134) where we also show (the line with crosses) the completion of f(x) determined analytically, by convolving h(x) and g(x) as given by (7.2) and (7.3) with the restriction $0 \leq x \leq 6$. From P15.1 we have $L_g = 6.0$.

Table 7.2 is a list of the various tests performed and the results obtained. As noted in the table, the completed f(x) obtained from procedure P15 compares favourably with that

Table 7.2

Summary of Tests on the Deconvolution of Incomplete Data

N_f+1	Noise added to f(x)[1]	M_G from P15.3	Zeros of G(w) from P15.2 & P15.3[2]	M_F (=$2m_f$) from P15.4	Completed f(x)	Filtering applied to F(u)/G(u) in P16.2[3]	Figure showing h(x) from P16.3
17	0	16	Shown in Fig. 7.3	32	Difference between theoretical completion of f(x) and its computed completion could not be seen in Fig. 7.1.	Cosine bell $u_{max} = .625$ / $u_d = .273$	7.5
21	0	20	Same as in Fig. 7.3 but with 2 more zeros.	40	Difference between theoretical completion of f(x) and its computed completion could not be seen in Fig. 7.1.	Cosine bell $u_{max} = .625$ / $u_d = .395$	Reconstructed h(x) is very similar to that in Fig. 7.5.
21	5%	20	Same as in Fig. 7.3 but with 2 more zeros.	40	Fig. 7.4.	Rectangular $u_{max} = .391$	7.6
35	5%	34	Same as in Fig. 7.3 but with 9 more zeros.	68	Similar to Fig. 7.4.	Cosine bell $u_{max} = .781$ / $u_d = .389$	7.7

1 The noise is normally distributed with standard deviation a fixed percentage of each datum.
2 The zeros ξ_G were computed by a procedure which takes advantage of the exponential form of g(x) (see section 8.4.2, p.167).
3 The cosine bell and rectangular filters are defined in section 8.1 (p.155).

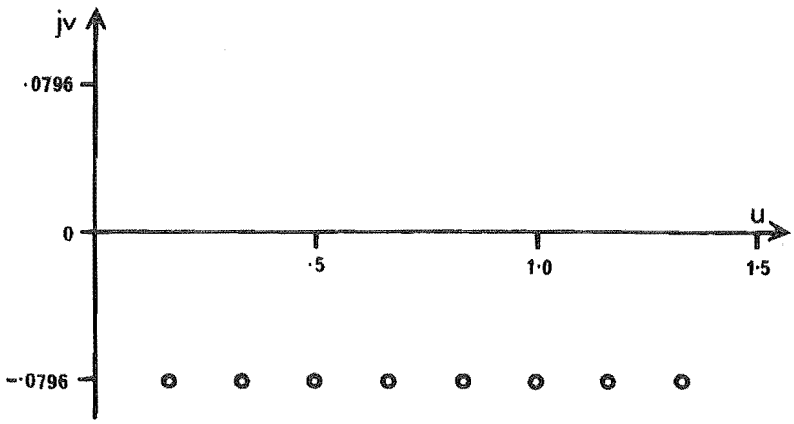


Fig. 7.3 The zeros of G(w) computed from the data (7.2) (p.133)

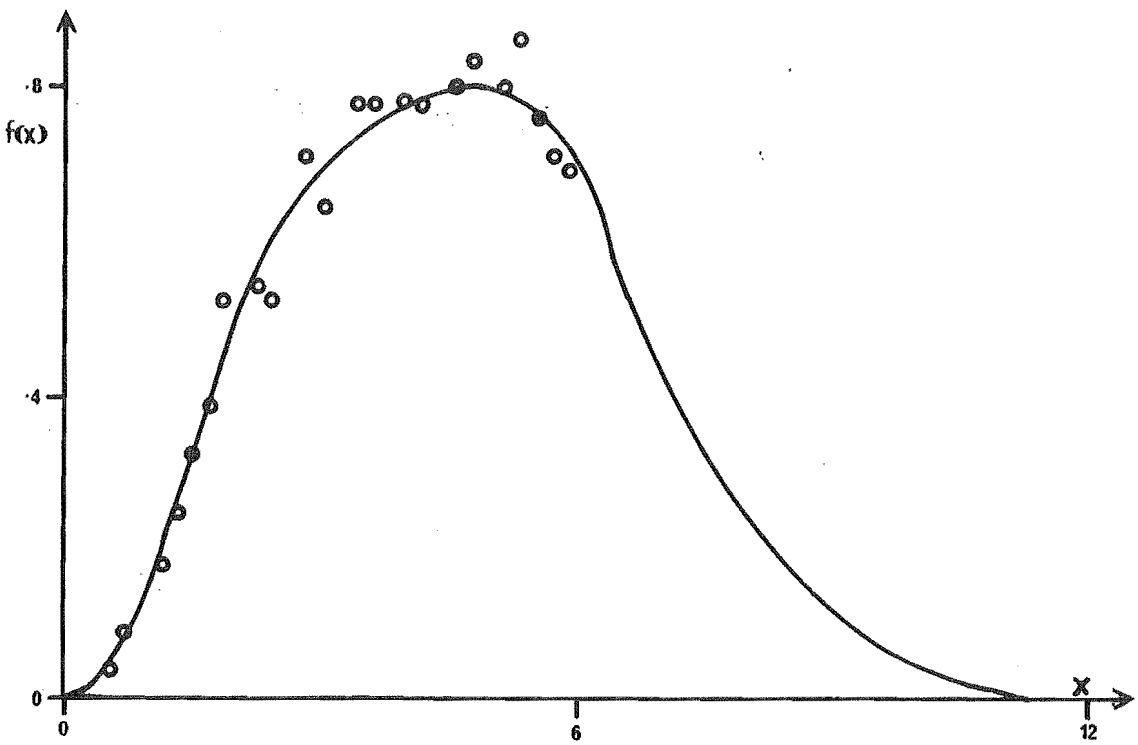


Fig. 7.4 The completed f(x) from data with 5% error for the case N_f = 20. —— completed function; OOO data.

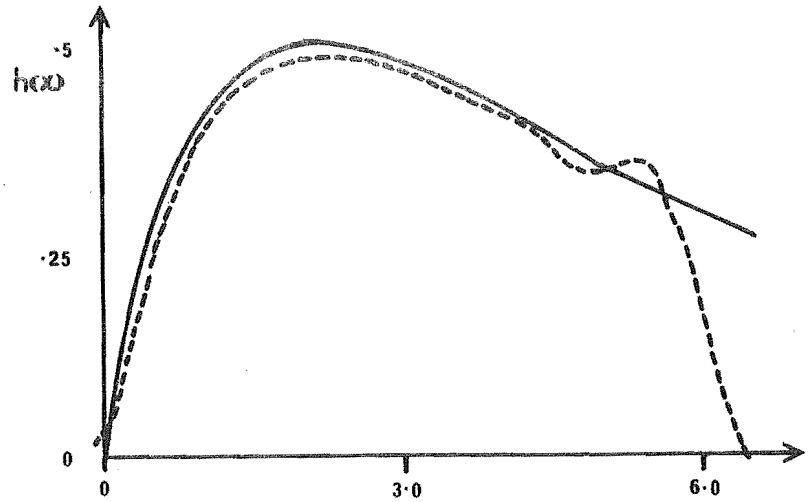


Fig. 7.5 ---- h(x) calculated from error-free data with $N_f = 16$. —— correct h(x) from equation (7.3) (p.133).

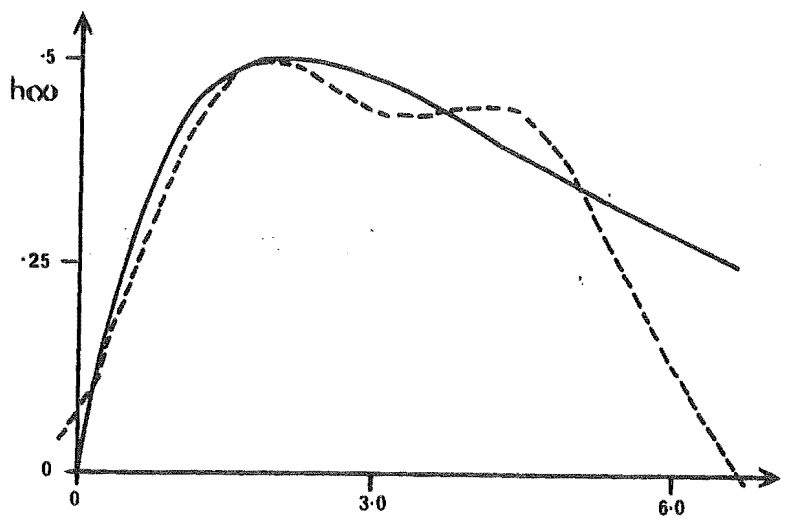


Fig. 7.6 ---- h(x) calculated from data with 5% error for the case $N_f = 20$. —— correct h(x) from equation (7.3) (p.133).

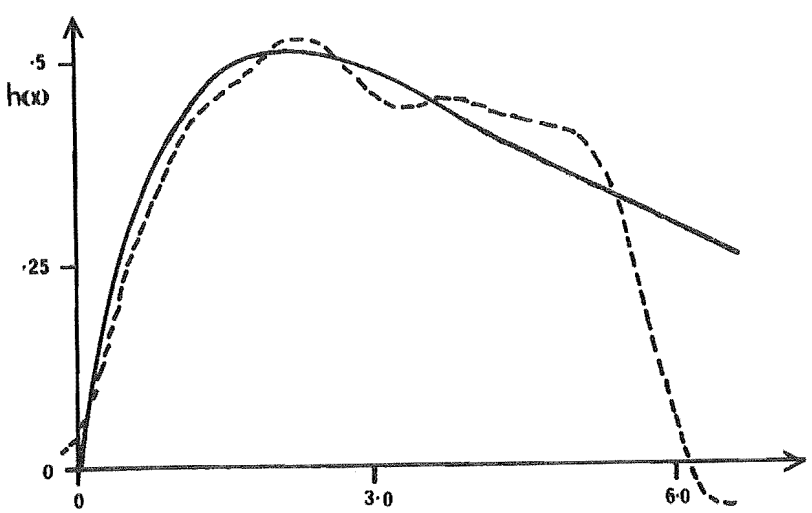


Fig. 7.7 ---- h(x) calculated from data with 5% error for the case $N_f = 34$. —— correct h(x) from equation (7.3) (p.133).

obtained analytically. However, to obtain acceptable results for the deconvolution it is necessary to filter the function F(u)/G(u). This is because m_f is finite, whereas we demonstrated in section 7.1 (p.138) that in principle it should be infinite. Notwithstanding, little change in h(x) was noticed as m_f was increased from 16 to 20. The tests in which noise was added to f(x) show the high sensitivity of deconvolution problems to error in the data. As with the error free results, there is little improvement with increasing m_f (cf. Fig. 7.6 for which $m_f = 20$, and Fig. 7.7 for which $m_f = 34$).

7.4 DECONVOLUTION OF CALCIUM ABSORPTION DATA

Deconvolution methods allow some clarification of the dynamics of the intestinal absorption of calcium and have been used by various authors (cf. Hart and Spencer, 1961; Silverman and Burgen, 1961; Birge et al., 1969; Szymendera et al., 1972). Two measurements using radioactive isotopes of calcium as tracers are conducted either simultaneously, using different isotopes (Ca^{45} and Ca^{47}), or on two separate occasions with the same isotope. If we assume that the stable calcium is at a steady state, the way in which tracer disappears from the blood is independent of the time of its introduction. Therefore, when the first tracer enters the blood after an oral dose, it will begin to decay away as if it were an intravenous injection of the same amount of tracer. The next small amount of tracer to appear in the blood will disappear similarly, but it will be delayed with respect to the first, and so on for all subsequent tracer from the oral dose.

The amount of orally ingested tracer in the blood at any one time f(t) is given by the convolution integral

$$f(t) = \int_0^t g(t-\tau)\ h(\tau)\ d\tau, \qquad (7.16)$$

where g(t) is the amount of tracer in the blood following an intravenous injection, and h(t) is the rate of 'initial entry' of tracer into the blood from the gut. By 'initial entry' we mean that tracer which has not been previously absorbed from the gut and then secreted.

In practice, the orally ingested tracer is given along with a known amount of stable calcium, so that the rate of initial entry of stable calcium into the blood can be found by dividing the h(t) determined from equation (7.16) by the specific activity of the original dose. However, if significant quantities of calcium are absorbed during the experiment, the stable calcium in the system cannot be at a steady state. Commonly, it is assumed that the absorbed calcium introduces negligible error.

The data shown as circles in Figures 7.8 and 7.9, made available by Dr J.H. Livesey of the Medical Unit, Princess Margaret Hospital, Christchurch, New Zealand, were obtained using the single isotope Ca^{47} in separate oral and intravenous experiments conducted six days apart. Clearly f(t) is incomplete. Because the data f(t) and g(t) are given at irregular intervals, it is necessary to interpolate them (we use the continuous lines which were drawn manually through the data in Figs 7.8 and 7.9) before applying procedure P15 (p.140). The data g(t) is further complicated by the absence of a point at t = 0. Such a measurement cannot be taken, as it coincides

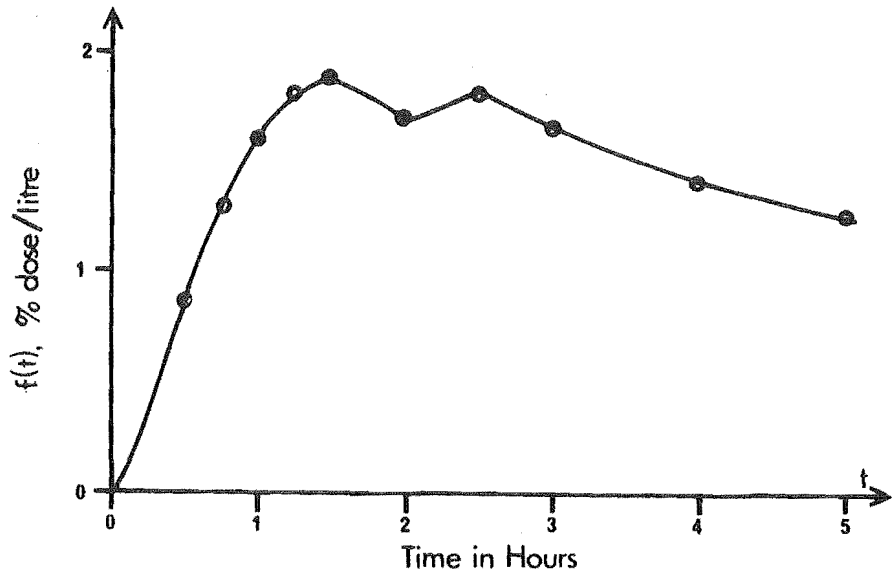


Fig. 7.8 Ca^{47} in the blood following an oral dose. OOOO data; —— interpolation.

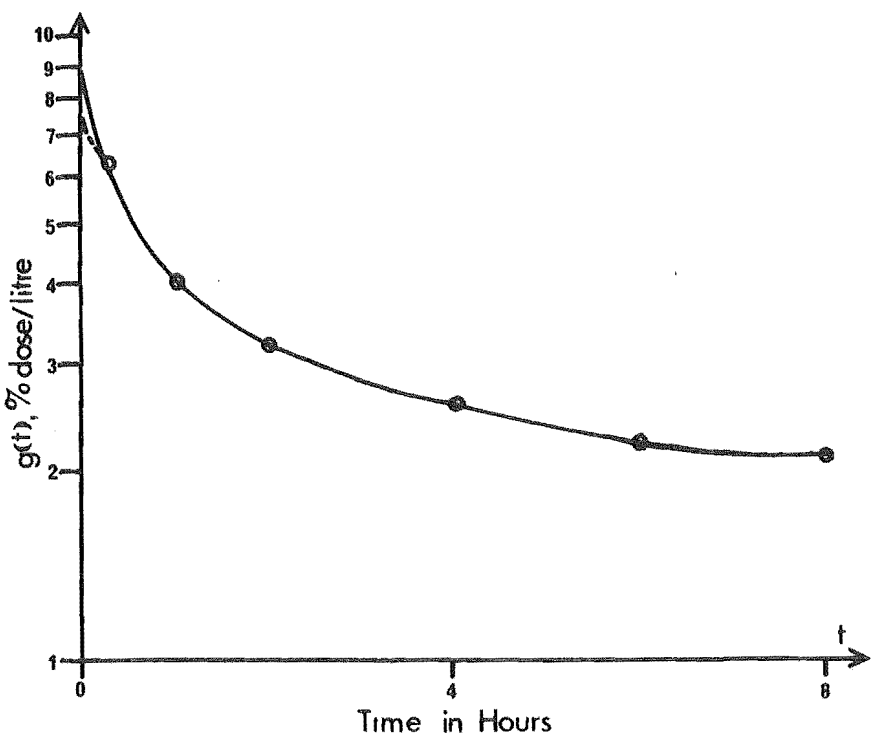


Fig. 7.9 Ca^{47} in the blood following an intravenous dose. OOOO data; —— interpolation; ---- alternative extrapolation to t = 0.

with the injection of the tracer. We extrapolate the data back to t = 0, and test the sensitivity of h(t) to the values obtained.

We choose $N_f = 34$ (our results varied little even with N_f as low as 20), so that from P15.1 (p.140) $L_g = 5.15$ hours. We note that although data for g(t) are given for t > 5.15 (cf. Fig. 7.9), our procedure does not use this information.

With the computer generated data in section 7.3, we found the zeros ξ_G using P15.3 after first multiplying g(x) by an increasing exponential function (see footnote 2 on p.142, and section 8.4.2, p.167). This operation gave no advantage with the present data in Fig. 7.9. The completed f(t) obtained from P15.4 and P15.5 (p.140) is shown in Fig. 7.10.

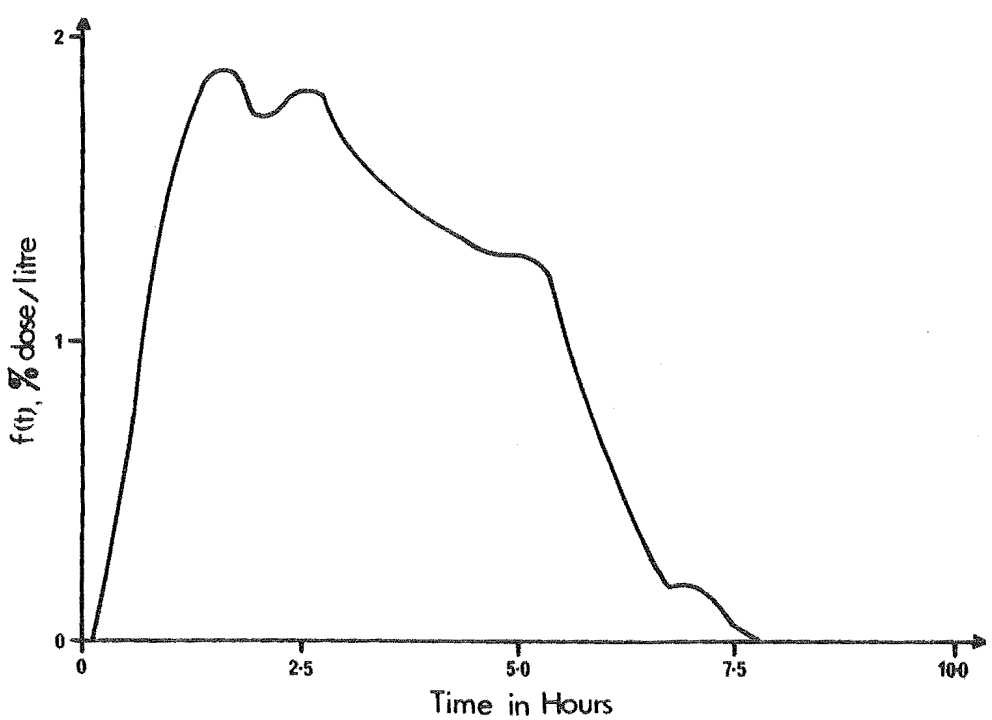


Fig. 7.10 The completed f(t) obtained from the calcium tracer data, Figs 7.8 and 7.9.

Although we have specified the completed f(t) to be of width 10.3 hours, it is small in the latter part of the interval. This suggests that h(t) is also small over a significant part of the interval $0 \leqslant t \leqslant 5.15$ hours. To deconvolve the completed data we plot $|F(u)/G(u)|$, which was obtained using P16.1 and P16.2 (p.141), in Fig. 7.11, and in Fig. 7.12 we show h(t) calculated using P16.3 after applying a rectangular window to F(u)/G(u) at both $u_{max} = 1.09$ hrs^{-1} and $u_{max} = .703$ hrs^{-1}. The choice of filter noticeably affects h(t), including its maximum value and the corresponding time.

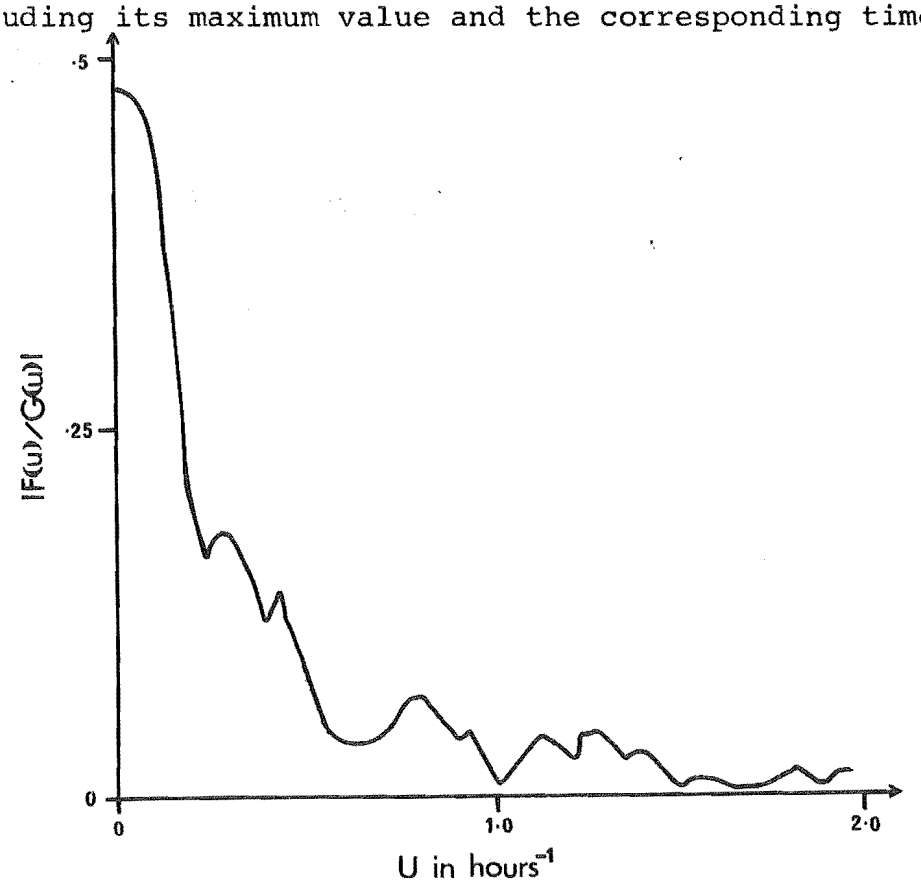


Fig. 7.11 $|F(u)/G(u)|$ obtained during the deconvolution of the calcium tracer data.

We cannot determine from the data whether the second, lower, peak in absorption rate actually occurs, or whether it is a consequence of experimental error, and should be filtered out.

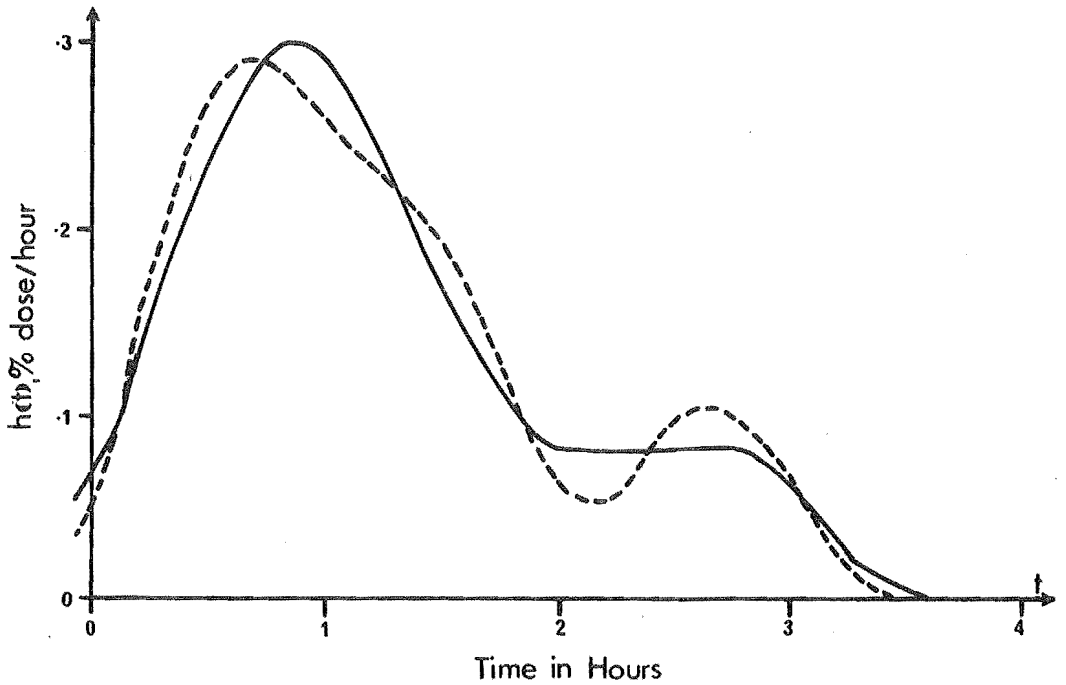


Fig. 7.12 Rate of initial entry of calcium from gut to blood. —— filter applied to $F(u)/G(u)$ at $u = .703$ hrs^{-1}; --- filter applied at $u = 1.09$ hrs^{-1}.

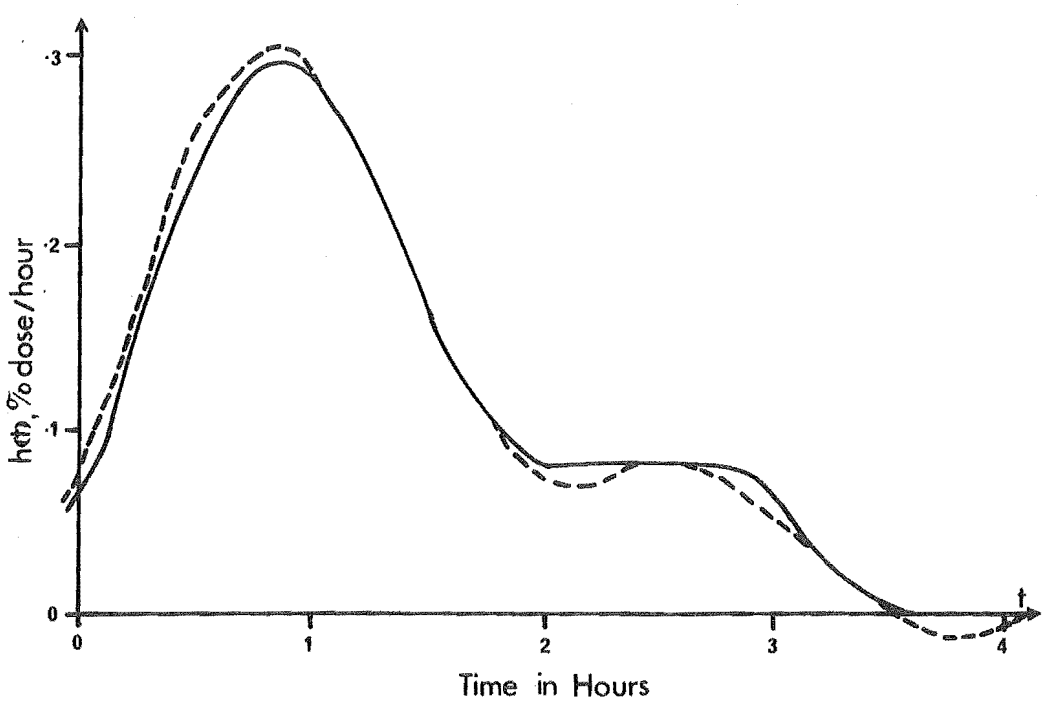


Fig. 7.13 The insensitivity of $h(t)$ to the extrapolation to $t = 0$ of the data $g(t)$ (Fig. 7.9, p.147). —— obtained from $g(0) = 9$. --- obtained from $g(0) = 7.5$.

That our deconvolution procedure is insensitive to the extrapolation of the g(t) data back to t = 0, is shown in Fig. 7.13, where we compare h(t) as obtained above, with that obtained using identical procedures, but with extrapolation of g(t) according to the dashed line in Fig. 7.9 (p.147).

Because we do not constrain h(t) to have any particular mathematical form, we cannot prevent it having significant value outside the range $0 \leqslant t \leqslant L_g$. In Figure 7.12 we see that the more F(u)/G(u) is filtered, the more h(t) takes on values for $t < 0$. Clearly h(t) should be zero for $t < 0$ because there is then no tracer in the system. The best filtering of F(u)/G(u) is a matter of judgement, and can be quickly determined by a series of trials.

When empirical models are used for the deconvolution of calcium tracer data, h(t) is constrained, perhaps incorrectly. Szymendera et al. (1972) represent the measured data by the models

$$f(t) = A_1 e^{-\lambda_1 t} - A_2 e^{-\lambda_2 t}, \qquad (7.17)$$

and

$$g(t) = A_3 e^{-\lambda_3 t} + A_4 e^{-\lambda_4 t}. \qquad (7.18)$$

They determine the parameters A_1, A_2, λ_1, λ_2, A_3, A_4, λ_3 and λ_4 by iteratively adjusting them to best fit the data (cf. section 3.3.1, p.54). Equations (7.17) and (7.18) are then substituted into the convolution integral (7.16) and h(t) is obtained by approximating (7.16) with a sum. However, their model (7.17) and (7.18) requires some comment.

Taking the Laplace transforms $F_\ell(s)$ of f(t), and $G_\ell(s)$ of g(t) using (3.59) (p.63) with appropriate changes in notation, yields

$$F_\ell(s) = \frac{A_1}{s+\lambda_1} - \frac{A_2}{s+\lambda_2}, \tag{7.19}$$

and

$$G_\ell(s) = \frac{A_3}{s+\lambda_3} + \frac{A_4}{s+\lambda_4}. \tag{7.20}$$

From (7.16) (p.146) and (3.59) (p.63) we have

$$F_\ell(s) = G_\ell(s)\, H_\ell(s) \tag{7.21}$$

so that

$$H_\ell(s) = \frac{[A_1(s+\lambda_2) - A_2(s+\lambda_1)](s+\lambda_3)(s+\lambda_4)}{[A_3(s+\lambda_4) + A_4(s+\lambda_3)](s+\lambda_2)(s+\lambda_1)}. \tag{7.22}$$

The initial value theorem (cf. Elgerd, 1967, p.552) states that

$$h(0) = \lim_{s\to\infty} s\, H_\ell(s), \tag{7.23}$$

and therefore, from (7.22) and (7.23)

$$h(0) \to \infty \tag{7.24}$$

unless

$$A_1 = A_2. \tag{7.25}$$

Clearly, (7.24) is unphysical and so we assume that (7.25) holds, although Szymendera et al. (1972) do not say whether they imposed this constraint in their procedure. Also, if (7.25) applies $H_\ell(s)$ can be expanded in partial fractions from which $h(t)$ can be determined as an analytical expression, thus avoiding the need to approximate the convolution integral with a sum.

CHAPTER 8

COMPUTATIONAL PROCEDURES

8.1 COMPUTATION OF FOURIER TRANSFORMS

The Fourier transform

$$F(w) = \int_{-\infty}^{\infty} f(x)\ e^{j2\pi wx}\ dx \qquad (8.1)$$

and its inverse

$$f(x) = \int_{-\infty}^{\infty} F(u)\ e^{-j2\pi ux}\ du \qquad (8.2)$$

are the basis of most of the techniques previously described. We calculate Fourier transforms both indirectly, when we first approximate the data by a set of functions the Fourier transform of which we know analytically (cf. the Hermite function representation (6.1), p.106 and its Fourier transform (6.2)), and directly, when we apply equation (8.1) to the data. In the former case, the numerical problem is one of approximating the data with the chosen set of functions, and we consider this in sections 8.3.2 and 8.4.1. For the latter case, we use the well-documented Fast Fourier Transform (FFT) algorithm (cf. Bergland, 1969). The FFT implementation used is called HARM (IBM, 1968).

The FFT algorithm actually computes a Fourier series but its use for the approximate calculation of the Fourier integral (8.1) when working with sampled data is well established (cf. Bergland, 1969). The accuracy of the approximation depends on the interval between samples, and the error introduced by replacing, with finite limits, the infinite range of integration in equations (8.1) and (8.2). Theoretically (cf.

Bergland, 1969), a function f(x) must be sampled at a rate equal to or greater than twice the maximum 'frequency' U_{max}, at which F(u) has significant value. Failure to satisfy this sampling condition results in 'aliasing', whereby the missing high frequency components are 'folded back' onto the lower frequency components, making them incorrect. For a sample interval T, the FFT computes F(u) at a maximum frequency u_{max}, given by

$$u_{max} = 1/2T. \tag{8.3}$$

For the computations in chapters 4, 5 and 6 the sampling of f(x) was found to be satisfactory provided $|F(u_{max})|$ was not greater than 1% of the maximum value of $|F(u)|$. Corresponding sampling criteria apply to F(u) when using the Fourier inverse formula (8.2) to obtain f(x).

We note in section 5.1 (p.93), that the examples are all of finite extent so the use of finite limits in (8.1) is exact. Procedures P7 and P8 in section 6.1.1 (p.112) require the Fourier transform of functions which are theoretically of infinite extent. However, in all cases described in sections 6.2 and 6.3 the data at the tails of the various f(x) and g(x) are given, as $|x|$ increases, until f(x) or g(x) falls to less than .5% of its maximum value. The error in using (8.1) with finite limits is therefore slight. However, the availability of data over only a finite interval when calculating the Fourier transform (4.37) (p.75) is one of the major limitations of the Mellin transform method for fitting exponential functions to data.

The most important source of error in the various Fourier transform calculations of previous chapters is due to the substitution of finite limits for the infinite integration in the Fourier inverse integral (8.2). In section 4.2.1 (p.77) we imposed finite limits because the error in and truncation of the data causes $R(\zeta)$ in equation (4.39) (p.75) to increase without bound with increasing ζ. Similarly $H(w)$ calculated during deconvolution by inverse filtering in various sections of chapters 5 and 7 diverges as u increases because of error in the data (see, for example, p.93). Clearly if $F(u)$ in (8.2) diverges as $|u|$ increases the integral does not exist. However, in all cases we experienced, the divergence is due to either the error in or truncation of the data. To obtain an inverse, we apply a filter (or 'window') to $F(u)$ to reduce the effects of the limitations in the data. We used two types of filter as indicated in Table 5.2 (p.94) and Table 7.2 (p.142). A rectangular window cuts $F(u)$ off sharply at a specified u_{max} such that

$$F(u) = 0 \qquad \text{for } |u| > u_{max}. \tag{8.4}$$

A rectangular window can produce unwanted oscillations in $f(x)$ which are reduced in amplitude by lowering $F(u)$ more gradually to zero, as does the cosine-bell window of width u_d:

$$F(u) = F(u_{max}) \{1 + \cos (\frac{u-u_{max}}{u_d})\pi\}$$

$$\text{for } u_{max} \leq u \leq u_{max}+u_d \tag{8.5}$$

and

$$F(u) = 0 \qquad \text{for } u \geq u_{max} + u_d. \tag{8.6}$$

There is a corresponding formula to reduce F(u) to zero in the interval $-u_{max} \geq u \geq -u_{max} - u_d$. A cosine bell window applied to R(ζ) (Fig. 4.3, p.77) remarkably reduces the amplitude of the auxilliary peaks in $Y(e^{-\beta})$ (cf. Fig. 4.4, p.78 and Fig. 4.2, p.76).

The FFT computes F(w) from (8.1) (p.153) for only real values of w. To calculate the Fourier transform for complex values of w which we require in section 5.1, we write

$$F(u+jv) = \int_{-\infty}^{\infty} f(x)\ e^{j2\pi(u+jv)x}\ dx \qquad (8.7)$$

so that

$$F(u+jv) = \int_{-\infty}^{\infty} f(x)\ e^{-2\pi vx}\ e^{j2\pi ux}\ dx. \qquad (8.8)$$

Therefore, to calculate F(u+jv) we first multiply f(x) by $e^{-2\pi vx}$ and then use the FFT. However, the sampling requirements for a successful FFT computation now apply to the modified data $f(x)\ e^{-2\pi vx}$. We found equation (8.8) satisfactory to calculate F(w) from f(x) given by equation (5.17) (p.93). On the other hand, the sampling requirements using (8.8) were found to be too stringent for g(x) given by (5.16) (p.93). (g(x) is a square pulse, which is more easily sampled than is $g(x)\ e^{-2\pi vx}$, which is an exponential function.) In this case, we calculate G(w) indirectly by first representing g(x) as a Fourier series similar to that in equation (7.4) (p.136)

$$g(x) = \sum_{m=-m_g}^{m_g} G_m\ e^{-j2\pi mx/L_g} \qquad 0 \leq x \leq L_g, \qquad (8.9)$$

and then evaluating its analytical Fourier transform

$$G(w) = e^{j\pi wL_g} \frac{\sin(\pi wL_g)}{\pi} \sum_{m=-m_g}^{m_g} \frac{G_m}{(w - \frac{m}{L_g})} . \quad (8.10)$$

After dividing the Fourier transforms F(w) and G(w) in the complex plane in section 5.1 (p.93), we calculate h(x) from H(w) using the following theory. If

$$h(x) = \int_{-\infty}^{\infty} H(u)\, e^{-j2\pi ux}\, du \quad (8.11)$$

and

$$h'(x) = \int_{-\infty}^{\infty} H(u+jv)\, e^{-j2\pi ux}\, du \quad (8.12)$$

then

$$h(x) = e^{2\pi vx}\, h'(x). \quad (8.13)$$

Therefore, to calculate h(x), we apply the FFT to H(u+jv) as if it were just H(u), but we multiply the resulting inverse by $e^{2\pi vx}$. In section 5.1 (p.95) we observed the magnification of errors in h(x) in the region where $e^{2\pi vx}$ is an increasing exponential function, but these errors lay outside the known extent of h(x) so we ignored them.

8.2 COMPUTATION OF THE GAMMA FUNCTION FOR COMPLEX ARGUMENT

In section 4.2 (p.74) we require the gamma function for complex argument for the analytic continuation of data onto the Bromwich contour using the Mellin transform. The gamma function is defined by (cf. Abramowitz and Stegun, 1965, p. 255)

$$\Gamma(z) = \int_0^{\infty} t^{z-1}\, e^{-t}\, dt \quad \mathrm{Re}\{z\} > 0, \quad (8.14)$$

and satisfies the recurrence relation

$$\Gamma(z+1) = z\Gamma(z). \tag{8.15}$$

The integral (8.14) does not provide a practical method for evaluating $\Gamma(z)$. In section 4.2.1 we used a double precision version of a program published by Lucas and Terrill (1971). They first augment z by an integer N so that $|z+N| > 10$ for which argument Stirling's asymptotic formula (cf. Abramowitz and Stegun, 1965, p.257)

$$\Gamma(z) \sim e^{-z} z^{z-\frac{1}{2}} (2\pi)^{\frac{1}{2}} \left[1 + \frac{1}{12z} + \frac{1}{288z^2} - \frac{139}{51840z^3} - \frac{571}{2488320z^4} + \ldots\right] \tag{8.16}$$

$$(z \to \infty \text{ in } |\arg z| < \pi)$$

is known to converge rapidly. They compute $\Gamma(z)$ from $\Gamma(z+N)$ by repeated application of the recurrence relation (8.15). Tests gave an accuracy of better than eight significant figures and a typical computation time of .05 seconds to evaluate $\Gamma(z)$.

8.3 COMPUTATIONS WITH HERMITE FUNCTIONS

8.3.1 Generation of Hermite Polynomials

In chapter 6 (p.106) we represent our 'complete' data by Hermite functions which are Hermite polynomials weighted with a Gaussian. The Hermite polynomial of order n is given by (cf. Abramowitz and Stegun, 1965, p. 775)

$$He_n(x) = n! \sum_{m=0}^{[\frac{n}{2}]} \frac{(-1)^m x^{n-2m}}{m!\, 2^m\, (n-2m)!}, \tag{8.17}$$

where $[\frac{n}{2}]$ denotes the largest integer not exceeding $\frac{n}{2}$. From (8.17) we see that

$$He_0(x) = 1 \text{ and } He_1(x) = x. \tag{8.18}$$

Instead of using (8.17) to directly evaluate the Hermite polynomials, we use the more efficient recurrence relation

$$He_{n+1}(x) = x\, He_n(x) - n\, He_{n-1}(x) \tag{8.19}$$

with (8.18) as starting values.

8.3.2 Fitting Hermite Functions to Data

To find the coefficients F_m in the Hermite function expansion (6.1) (p.106), we use the formula (6.16) (p.109) which is derived using the orthogonal properties of Hermite polynomials (cf. equation (6.14)). We are given only samples of $f(x)$ and therefore we must use an approximate formula for the integration in (6.16) as follows.

1. Multiply the data $f(x)$ by $e^{-x^2/a_f^2} He_n(2x/a_f)$.
2. Integrate $f(x)e^{-x^2/a_f^2} He_n(2x/a_f)$ using Simpson's rule (cf. Abramowitz and Stegun, 1965, p.886).
3. Interpolate midway between each adjacent pair of samples of $f(x)$ and again multiply by $e^{-x^2/a_f^2} He_n(2x/a_f)$.
4. Integrate the interpolated samples using Simpson's rule.
5. Repeat the interpolation 3 and integration 4 until successive integrals agree to within three significant figures.

By this procedure we ensure that there are sufficient points in the integration to maintain accuracy. The choice of interpolation formula in step 3 is arbitrary, but we used cubic spline functions (cf. Greville, 1967). Cubic splines interpolate with a cubic polynomial between given samples or 'knots', with the constraint that the first order derivatives of

adjacent cubic polynomials are equated at each knot. Interpolation with cubic splines is similar to manual interpolation on graph paper. After completion of the computations described in chapter 6, it was realised that the M_F+1 numerical integrations of equation (6.16) (p.109) needed to find M_F+1 coefficients in equation (6.1) (p.106), could by replaced by the inversion of a single matrix of order M_F+1. The theory is given by Lanczos (1956, p.371). In principle, the exact representation of a function f(x) by a series of the form (6.1) requires an infinite number of terms, so that, ideally

$$f(x) = e^{-x^2/a_f^2} \sum_{m=0}^{\infty} F_m \, He_m(2x/a_f). \qquad (8.20)$$

If we truncate the series at $m = M_F$, the remainder η_{M_F} is given by

$$\eta_{M_F}(x) = e^{-x^2/a_f^2} \sum_{M_F+1}^{\infty} F_m \, He_m(2x/a_f). \qquad (8.21)$$

Assuming that the series (8.20) has quick convergence, we may estimate the remainder of $\eta_{M_F}(x)$ by keeping only the first term in (8.21) so that

$$\eta_{M_F}(x) = e^{-x^2/a_f^2} F_{M_F+1} \, He_{M_F+1}(2x/a_f). \qquad (8.22)$$

$\eta_{M_F}(x)$ is zero at the zeros (or roots) of the polynomial $He_{M_F+1}(2x/a_f)$, so that we can obtain the coefficients of the finite expansion

$$f_{M_F+1}(x) = e^{-x^2/a_f^2} \sum_{m=0}^{M_F} F_m' \, He_m(2x/a_f) \qquad (8.23)$$

by fitting the functional values f(x) at the M_F+1 zeros β_i of

$He_{M_F+1}(2x/a_f)$:

$$\sum_{m=0}^{M_F} F_m' \, He_m(2\beta_i/a_f) = f(\beta_i) \qquad i = 1,2,\ldots,M_F+1 \quad (8.24)$$

such that

$$He_{M_F+1}(2\beta_i/a_f) = 0 \qquad i = 1,2,\ldots,M_F+1. \quad (8.25)$$

That M_F+1 zeros of $He_m(2x/a_f)$ do exist for real values of x is well known (cf. Lanczos, 1956, p.373). Although the solutions F_m' of the linear algebraic equations (8.24) will not normally coincide with the coefficients F_m obtained by integration of equation (6.16) (p.109), Lanczos (1956, p.372) suggests that the error will not be essentially worse, but the computational savings will be great. Also, the procedure of equation (8.24) provides an alternative method for estimating M_F (cf. section 6.1.2). The data f(x) falls asymptotically to zero as x increases (cf. Fig. 6.1, p.105), and so we can identify x_{max}, the value of x beyond which the data f(x) cannot be distinguished from the experimental error. Because the furthest out zero of $He_n(2x/a_f)$ moves away from the origin as n increases, we can fix M_F+2 as the value of n for which the furthest out zero lies at $x > x_{max}$.

8.3.3 Optimisation of a_f

Procedure P4 (p.111) for a fixed M_F or M_G, requires the iterative adjustment of a_f or a_g to minimise the RMS error (6.17) (p.109) between the data and its representation (6.1) (p.106) by Hermite functions. a_f is nonlinearly related to f(x) in equation (6.1) so that its adjustment is a nonlinear optimisation problem. The Fibonacci search (cf. Dixon, 1972, p.20) is ideally suited to nonlinear optimisation with respect

to one parameter (in our case a_f or a_g). Provided $\frac{\partial R}{\partial a_f}$ (R is the RMS error) is zero for only one value of a_f in the stated range over which searching is to be carried out (i.e. provided there is only one 'minimum'), the Fibonacci search systematically finds the optimum a_f using a minimum number of trial values. Typically, to find the optimum a_f in the range $4 \leqslant a_f \leqslant 11$ to an accuracy of .01 requires R to be computed at only 15 test values of a_f.

8.3.4 Calculation of Zeros and Reconstruction

When data f(x) is expanded in terms of Hermite functions (6.1) (p.106) the zeros of F(w) are shown in section 6.1 (p.106) to be the zeros of the polynomial $P_{M_F}(w)$ given by

$$P_{M_F}(w) = a_f\sqrt{\pi}/A_F \sum_{m=0}^{M_F} j^m F_m He_m(2\pi a_f w). \qquad (8.26)$$

where A_F is given by equation (6.4) (p.106). Substituting for the Hermite polynomials from (8.17) (p.158) we have

$$P_{M_F}(w) = a_f\sqrt{\pi}/A_F \sum_{m=0}^{M_F} j^m F_m \sum_{i=0}^{[\frac{m}{2}]} \frac{(-1)^i (2\pi a_f)^{m-2i} w^{m-2i}}{i!\, 2^i\, (m-2i)!} \qquad (8.27)$$

Equation (8.27) is a sum of M_F+1 polynomials of increasing order. We calculate the coefficients of w^n separately for each polynomial, and then add them.

To reconstruct f(x) from the M_F zeros $w_{F,m}$ of F(w) (procedure P2, p.107), we let

$$f(x) = e^{-x^2/a_f^2} Q_{M_F}(x), \qquad (8.28)$$

where $Q_{M_F}(x)$ is a polynomial in x of order M_F. If

$$f_0(x) = \int_{-\infty}^{\infty} F_0(u)\ e^{-j2\pi ux}\ du \tag{8.29}$$

(the inverse Fourier transform), and

$$f_1(x) = \int_{-\infty}^{\infty} uF_0(u)\ e^{-j2\pi ux}\ du, \tag{8.30}$$

then,

$$f_1(x) = f_o'(x)/(-j2\pi), \tag{8.31}$$

where $f_0'(x)$ denotes $df_0(x)/dx$.

Now from equations (6.3) and (6.6) (p.106)

$$F(u) = A_F\ e^{-\pi^2 a_f^2 u^2} \prod_{m=1}^{M_F} (u-w_{F,m}) \tag{8.32}$$

so we define

$$F_0(u) = A_F\ e^{-\pi^2 a_f^2 u^2} \tag{8.33}$$

and

$$F_n(u) = F_{n-1}(u)\ (u-w_{F,n}) \qquad 1 \leqslant n \leqslant M_F. \tag{8.34}$$

From equations (8.29)-(8.34) we deduce the recurrence formula

$$f_n(x) = f'_{n-1}(x)/(-j2\pi) - w_{F,n} f_{n-1}(x) \qquad 1 \leqslant n \leqslant M_F \tag{8.35}$$

with the starting formula

$$f_0(x) = \frac{A_F}{a_f\sqrt{\pi}}\ e^{-x^2/a_f^2} \tag{8.36}$$

derived from equations (8.29) and (8.33). At any stage in the recurrence scheme (8.35), $f_{n-1}(x)$ is a polynomial in x multiplied by a Gaussian, and so $f'_{n-1}(x)/(-j2\pi)$ is of the same form and is easily added to $w_{F,n} f_{n-1}(x)$.

In the various reconstructions reported in chapter 6, we set the term A_F (or its equivalent A_G or A_H) in equation (8.33) to unity, as we were interested in only a normalised f(x).

8.4 COMPUTATIONS WITH FOURIER SERIES

8.4.1 Fitting Fourier Series to Data

The fast Fourier transform algorithm calculates the Fourier series coefficients F_m in (7.4) (p.136) from data f(x) very rapidly. However, the particular implementation HARM (IBM, 1968) requires 2^N (N an integer) samples of f(x). This restriction makes HARM unsuitable for some of the computations reported in chapter 7 where we have, for example 21 sample points (cf. Table 7.2, p.142). We found it more convenient (and not much less efficient) to directly solve (7.4) as the set of linear algebraic equations

$$\sum_{m=-m_f}^{m_f} F_m \, e^{-j2\pi m x_n / L_f} = f(x_n) \qquad 0 \leq n \leq N_f \tag{8.37}$$

where the $f(x_n)$ are the data. Because f(x) is real

$$F_m = F^*_{-m}, \tag{8.38}$$

and so we may halve the required computer storage by solving for the coefficients $A_m = \mathrm{Re}\{F_m\}$ and $B_m = \mathrm{Im}\{F_m\}$, the real and imaginary parts of F_m, in

$$F_0 + 2 \sum_{m=1}^{m_f} A_m \cos(2\pi m x_n / L_f) + B_m \sin(2\pi m x_n / L_f) = f(x_n). \qquad 0 \leq n \leq N_f \tag{8.39}$$

To solve (8.39) for 35 coefficients F_m takes, typically, 6 seconds using the method given in section 8.6. When we complete the data f(x) using procedure P15 (p.140) we must solve the combined equations (7.14) and (7.15) (p.140) for F_m. We rewrite the M_G equations (7.14) as

$$(e^{j2\pi w_{G,n} L_f} - 1) \sum_{m=-m_f}^{m_f} \frac{F_m}{j2\pi(w_{G,n} - \frac{m}{L_f})} = 0$$
$$1 \leqslant n \leqslant M_G, \qquad (8.40)$$

so that from (8.38)

$$F_0 \frac{e^{j2\pi w_{G,n}} - 1}{j2\pi w_{G,n}} + \sum_{m=1}^{m_f} (A_m + jB_m) \left\{ \frac{e^{j2\pi w_{G,n}} - 1}{j2\pi(w_{G,n} - \frac{m}{L_f})} \right\}$$
$$+ (A_m - jB_m) \left\{ \frac{e^{j2\pi w_{G,n} L_f} - 1}{j2\pi(w_{G,n} + \frac{m}{L_f})} \right\} = 0 \qquad (8.41)$$

$$\text{for} \quad 1 \leqslant n \leqslant M_G$$

Because we have imposed the constraint (8.38), we need only include the $m_g = M_G/2$ zeros of G(w) which are either in the left half or the right half of the complex w-plane. We define

$$S_n = \frac{e^{j2\pi w_{G,n} L_f} - 1}{j2\pi}, \qquad (8.42)$$

$$C_n = S_n / w_{G,n}, \qquad (8.43)$$

$$D_{m,n} = \frac{S_n}{(w_{G,n} - \frac{m}{L_f})}, \qquad (8.44)$$

and

$$E_{m,n} = \frac{S_n}{(w_{G,n} + \frac{m}{L_f})}, \qquad (8.45)$$

so that, since the real and imaginary parts of F(w) in (8.41) are zero,

$$F_0 \mathrm{Re}\{C_n\} + \sum_{m=1}^{m_f} A_m \mathrm{Re}\{D_{m,n} + E_{m,n}\}$$

$$+B_m \mathrm{Im}\{E_{m,n} - D_{m,n}\} = 0 \qquad 1 \leq n \leq M_G \tag{8.46}$$

and

$$F_0 \mathrm{Im}\{C_n\} + \sum_{m=1}^{m_f} A_m \mathrm{Im}\{D_{m,n} + E_{m,n}\}$$

$$+ B_m \mathrm{Re}\{D_{m,n} - E_{m,n}\} = 0 \qquad 1 \leq n \leq M_G. \tag{8.47}$$

Expressed in the form (8.47), (8.46) and (8.39), the combined equations (7.14) and (7.15) are solved for A_m and B_m as indicated in section 8.6.

8.4.2 Calculation of Zeros

Procedure P13 (p.137) for finding the zeros of $F(w)$ requires the calculation of the complex coefficients γ_m in the polynomial $P_F(w)$ (7.8) (p.136) such that

$$P_F(w) = \sum_{m=0}^{M_F} \gamma_m w^m. \tag{8.48}$$

We expand (7.8) to give

$$P_F(w) = F(0) \prod_{m=1}^{m_f} \left(w^2 - \frac{m^2}{L_f^2}\right) + \sum_{n=1}^{m_f} \prod_{m=1}^{m_f(n)} \left(w^2 - \frac{m^2}{L_f^2}\right)$$

$$\left\{w^2 \left(F\left(\frac{n}{L_f}\right) + F\left(\frac{-n}{L_f}\right)\right) + \frac{wn}{L_f} \left(F\left(\frac{n}{L_f}\right) - F\left(\frac{-n}{L_f}\right)\right)\right\}. \tag{8.49}$$

Equation (8.49) consists of m_f+1 individual polynomials, the corresponding coefficients of which are added to give the γ_m in (8.48). The zeros of $P_F(w)$ are calculated from (8.48) using the method of section 8.5.

For the tests with computer-generated data in section 7.3, g(x) (7.2) (p.133) decays exponentially. The Fourier series representation of a decaying exponential function does not converge rapidly, and so, where indicated in section 7.3 we employed the following device to improve the accuracy of the zeros of G(w). An estimate of the exponential decay constant α_d for the data g(x) was obtained, usually by a graphical form of the peeling method (cf. section 3.3.1, p.52) and the data were then multiplied by the increasing exponential with the same exponent. Equation (8.8) (p.156) indicates that first multiplying g(x) by $e^{\alpha_d x}$ has the effect on G(w) of shifting the real line down to the coordinate $V = -\frac{\alpha_d}{2\pi}$ in the complex w-plane. We therefore represent g(x) $e^{\alpha_d x}$ by a Fourier series, calculate the zeros of its Fourier transform, and finally subtract $\frac{\alpha_d}{2\pi}$ from the imaginary part of each zero to compensate for the shifting of the real line.

8.5 ZEROS OF COMPLEX POLYNOMIALS

To calculate the zeros of the complex polynomials (6.5) (p.106) and (7.8) (p.136) we use the program CPOLY published by Jenkins and Traub (1972). Their procedure, which calculates in double precision arithmetic, is both accurate and efficient. It computes, for example, the zeros of a polynomial of order 32 in approximately 15 seconds. The mathematical details of the algorithm used in CPOLY are given in Jenkins and Traub (1970). The program POLRT (IBM, 1968) as converted by Napier (1972) for use with complex polynomials, was also tried, but in some instances the dynamic range of the computer (10^{-72} to 10^{72}) was exceeded during the computations, giving incorrect

results. In CPOLY all quantities are scaled to ensure that numbers do not go outside the permitted range.

8.6 SOLUTION OF LINEAR ALGEBRAIC EQUATIONS

To solve the linear algebraic equations (4.55) (p.85) for Y_m we used the Gauss elimination subroutine DGELG (IBM, 1968) which computes in double precision. In section 4.3 (p.86) we indicated that the computed Y_m values were such as to reconstruct $y(t_n)$ accurate to seven significant figures when substituted back into equation (4.55). However, this result is no guarantee of the accuracy of the Y_m values, as the following example, taken from Westlake (1968, p.89) shows.

The correct solution to

$$x_1 + 10\,x_2 = 11 \tag{8.50}$$

and

$$10x_1 + 101x_2 = 111 \tag{8.51}$$

is

$$x_1 = x_2 = 1.$$

For any x_1 and x_2 we define the residuals r_1 and r_2 by

$$r_1 = (11 - x_1 - 10x_2)^2 \tag{8.52}$$

and

$$r_2 = (111 - 10x_1 - 101x_2)^2 \tag{8.53}$$

so that if $x_1 = x_2 = 1$ then $r_1 = r_2 = 0$. However, if $x_1 = 1.001$ and $x_2 = 1.01$, $r_1+r_2 = 1.05$, whereas if $x_1 = 11.1$ and $x_2 = 0$, $r_1+r_2 = .01$ which is less, even though the values of x_1 and x_2 are further from the correct solution.

In contrast, the various equations (8.39), (8.46) and (8.47) (p.164-66) for calculating the coefficients of the Fourier series expansion (7.4) (p.136) of f(x) (or g(x)) are well conditioned, the results being insensitive to small errors in the data. We obtained satisfactory solutions using the double precision subroutine DLLSQ (IBM, 1968). DLLSQ is intended to solve over-determined linear algebraic equations (i.e. when there are more equations than unknowns) in a least squared sense. It was used in the computations of chapter 7 in case attempts were to be made to 'smooth' the data by using more data points than Fourier coefficients. However, this approach was not pursued.

CHAPTER 9

MODELLING OF GASTRO-INTESTINAL FLUID FLOW

Unlike the models listed in Table 2.1 (p.19), CAMET2 incorporates a description of gastro-intestinal fluid flow. Worthwhile conclusions about the value of some clinical tests for calcium loss to the intestine have resulted (see simulations 7 and 8, Table 2.5, p.34). The distal flow of calcium in CAMET2 depends directly on its concentration, and on the rate of fluid flow (eqns. (2.24) and (2.35), p.26). Therefore, the availability of calcium for absorption is determined by the movement of material in the intestine, which is taken to be independent of the calcium present.

9.1 THE INCONSISTENCY OF GUT FLUID FLOW IN CAMET2

In CAMET2 the gut is divided into six compartments of equal length, each one containing calcium and fluid. The rate of flow of fluid from one compartment to the next is given by (cf. equations (2.31) and (2.32), p.27)

$$RF_{I(I+1)} = RF_{(I-1),I}(1-AF_I) + RF_{1,I} \quad 6 \leqslant I \leqslant 10 \qquad (9.1)$$

and

$$RF_{fecl} = RF_{10,11}(1 - AF_{11}) + RF_{1,11} \qquad (9.2)$$

where AF_I is a constant which represents the fraction of fluid which is absorbed, and $RF_{1,I}$ represents the rate of secretion into the Ith compartment. We noted in section 2.2 (p.30) that (9.1) and (9.2) imply that the volume of a compartment V_I does not change. However, a conflict arises, because in CAMET2 V_I

is calculated from the varying rate $RF_{I(I+1)}$ according to (cf. equations (2.33) and (2.34), p.27)

$$V_I = B_{I,1} RF_{I(I+1)} / (B_{I,2} + RF_{I,(I+1)}), \quad (9.3)$$

and

$$V_{11} = B_{11,1} RF_{fec1} / (B_{11,2} + RF_{fec1}). \quad (9.4)$$

When a meal enters the stomach, the fluid flow $RF_{5,6}$ from stomach to duodenum increases, but because of (9.1) and (9.2) all the other distal fluid flows rise at the same instant. There follows an immediate increase in the volume of the compartments according to (9.3) and (9.4), and a consequent drop in calcium concentration.

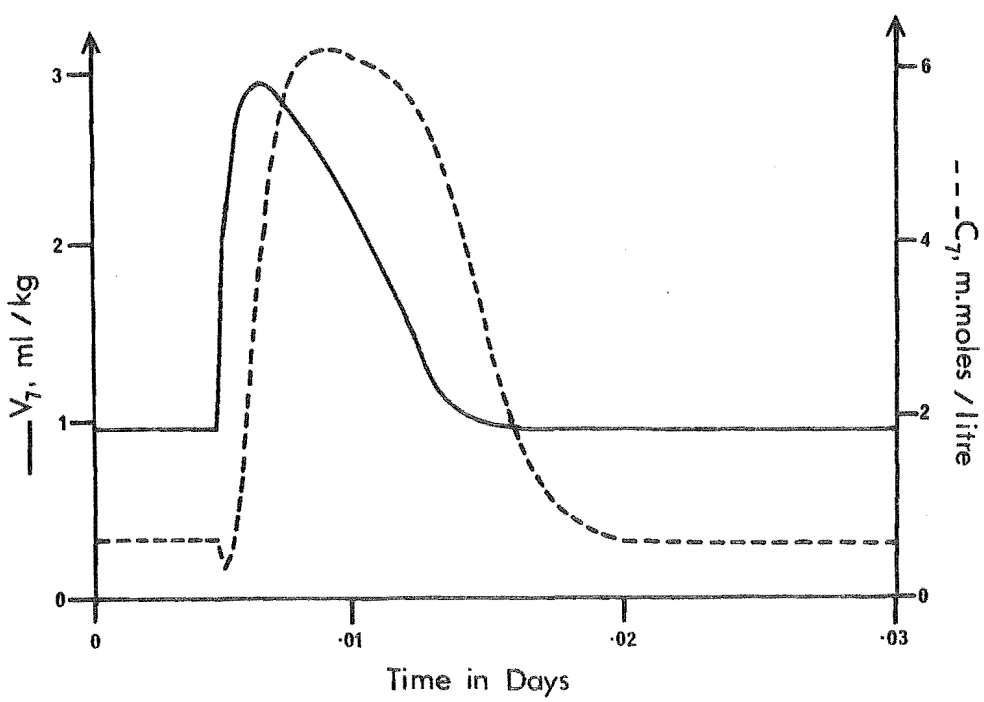


Fig. 9.1 The effect of the inconsistent fluid equations in CAMET2 on the volume and calcium concentration of compartment 7.

We see the quantitative effect in Fig. 9.1 where a 22 min.

meal consisting of 7.7 ml/kg of fluid and .083 m.moles/kg of calcium has been simulated using the particular CAMET2 equations listed in Table 9.2, and the values of the various parameters listed in Table 2.4 (p.31). (Unless otherwise stated all simulations reported in this chapter were carried out using the program SIMUL8 which is described in chapter 11). At the onset of the meal the volume of compartment 7 immediately increases, and its calcium concentration falls, recovering only when the calcium taken with the meal arrives.

The relationships (9.1)-(9.4), which were introduced into CAMET by Livesey (1970), are based on measurements of liquid volume and flow rate in various sections of the intestine, by Dillard et al. (1965). Essentially, the volume increased linearly with flow rate for flows less than 5 ml/min., but for higher flows, the volume exhibited saturation. The volumes of segments of the jejunum were approximately 1.6 times the volumes of segments of the ileum. The parameters AF_I, originally derived by Livesey (1970), and later modified by Pearson (1972), were obtained by considering average figures for human fluid intake, absorption, secretion, and loss in the faeces.

We modify the gut fluid equations to give CAMET3 as in Table 9.3. The relationship (9.3) (p.171) between fluid flow and volume is maintained, but in the rearranged form (9.28). We note that the rate at which material leaves the stomach, given by equation (9.27), is a modification of the CAMET2 expression (9.12). The reason for this is given in section 9.2 (p.181). We introduce the differential $\dot{V}_I$ - the difference between the fluid flows into and out of a compartment.

Table 9.2

CAMET2 Equations Used to Show the Inconsistency of Gut Fluid Flow

Differential Equations

Stomach

$$\dot{V}_5 = RF_{diet} + RF_{1,5} - RF_{5,6} \quad (9.5)$$

$$\dot{M}_5 = R_{diet} + R_{1,5} - R_{5,6} \quad (9.6)$$

Intestine

$$\dot{M}_I = R_{(I-1),I} + R_{1,I} - R_{I,(I+1)} - R_{I,1} \qquad 6 \leq I \leq 10 \quad (9.10)$$

$$\dot{M}_{11} = R_{10,11} + R_{1,11} - R_{fecl} - R_{11,1} \quad (9.11)$$

Rate Expressions and Volumes

Stomach

$$*RF_{5,6} = F\, V_5^{\frac{1}{2}} \quad (9.12)$$

$$R_{5,6} = RF_{5,6} C_5 \quad (9.13)$$

$$**R_{1,5} = K_c\, RG_5 \quad (9.14)$$

Intestine

$$RF_{1,6} = 3RF_{5,6} + G_1 \quad (9.15)$$

$$R_{1,6} = K_c(G_2 RF_{1,6} + G_3) \quad (9.16)$$

$$R_{1,I} = K_c\, RG_I \qquad 7 \leq I \leq 11$$

$$*RF_{I,(I+1)} = RF_{(I-1),I}(1-AF_I) + RF_{1,I} \qquad 6 \leq I \leq 10 \quad (9.17)$$

$$*RF_{fecl} = RF_{10,11}(1 - AF_{11}) + RF_{1,11} \quad (9.18)$$

$$*V_I = B_{I,1} RF_{I,(I+1)} / (B_{I,2} + RF_{I,(I+1)}) \qquad 6 \leq I \leq 10 \quad (9.19)$$

$$*V_{11} = B_{11,1} RF_{fec1} / (B_{11,2} + RF_{fec1}) \quad (9.20)$$

$$R_{I,(I+1)} = RF_{I,(I+1)} \; C_I \qquad 6 \leqslant I \leqslant 10 \quad (9.21)$$

$$R_{fec1} = RF_{fec1} \; C_{11} \quad (9.22)$$

$$R_{I,1} = E_{I,1} C_I / (E_{I,2} + C_I) \qquad 6 \leqslant I \leqslant 11 \quad (9.23)$$

Concentrations

$$C_I = M_I / V_I \qquad 5 \leqslant I \leqslant 11 \quad (9.24)$$

* These equations are modified or replaced in CAMET3 (Table 9.3, below).

** We set $K_c = 1$ because we are not considering fluctuations in the concentration of calcium in the blood.

Table 9.3

CAMET3 Fluid Equations

New Differential Equations

$$\dot{V}_I = RF_{(I-1),I} + RF_{1,I} - RF_{I,(I+1)} - RF_{I,1} \qquad 6 \leqslant I < 10 \quad (9.25)$$

$$\dot{V}_{11} = RF_{10,11} + RF_{1,11} - RF_{fec1} - RF_{11,1} \quad (9.26)$$

Fluid Flow Rates

$$*RF_{5,6} = F \; V_5 (AF_5 - V_6) \quad (9.27)$$

$$*RF_{I,(I+1)} = V_I \; B_{I,2} / (B_{I,1} - V_I) \qquad 6 \leqslant I \leqslant 10 \quad (9.28)$$

$$*RF_{fec1} = V_{11} \; B_{11,2} / (B_{11,1} - V_{11}) \quad (9.29)$$

* denotes a modification to existing CAMET2 equations (cf. Table 9.2, p.173).

By simulating CAMET3 for 0.5 days with no intakes, we arrive at the initial conditions for the various quantities in the gut which are shown, along with any new or changed model parameters, in Table 9.4.

A repeat of the previous simulation consisting of a 22 minute meal of 7.7 ml/kg of fluid and .083 m.moles/kg of calcium was carried out using the equations of Table 9.2 modified as in Table 9.3. Now the volume of compartment 7 increases gradually (Fig. 9.5) and its calcium concentration, also shown in Fig. 9.5, does not drop at the beginning of the meal (cf. Fig. 9.1, p.171). The calcium is not retained as long in compartment 7 of CAMET3.

However, the modified equations in Table 9.3 have little overall effect on the model predictions. Some results of a simulation of a three meal day are given in Table 9.6. The meals, each of 22 minutes duration, occur at 8 a.m., 1 p.m. and 6 p.m., and consist of .083 m.moles/kg (216 mg) of calcium, and 7.7 ml/kg (500 ml) of fluid. We note that the accumulated faecal calcium of 380 mg from CAMET2 is greater by 34 mg than that obtained in a similar simulation by Pearson (1973). The difference occurs because in our model (Table 9.2, p.173), we have constrained the blood calcium to remain at a constant level. Bearing in mind the large variation that occurs in nature, the modified fluid equations adequately represent the given data. Because, in (9.28), we have not altered the relationship (9.18) between volume and flow derived by Livesey (1970), CAMET3 must still describe the data of Dillard et al. (1965).

Table 9.4

CAMET3 Initial Conditions and Model Parameters

Parameter	Value	Units
Differential Equation Initial Conditions		
M_1	.45	m.moles/kg
M_2	.52	"
M_3	1.23	"
M_4	.87	"
M_5	$.18 \times 10^{-3}$	"
M_6	$.11 \times 10^{-2}$	"
M_7	$.65 \times 10^{-3}$	"
M_8	$.21 \times 10^{-3}$	"
M_9	$.11 \times 10^{-3}$	"
M_{10}	$.65 \times 10^{-4}$	"
M_{11}	$.4 \times 10^{-4}$	"
MT_1 to MT_{11}	0	"
M_c	7.1	ng/kg
M_p	93.0	"
M_{bca}	.4242	moles/kg
M_{bcoll}	21.23	"
V_5	$.118 \times 10^{-3}$	litres/kg
V_6	$.128 \times 10^{-2}$	"
V_7	$.982 \times 10^{-3}$	"
V_8	$.629 \times 10^{-3}$	"
V_9	$.294 \times 10^{-3}$	"
V_{10}	$.205 \times 10^{-3}$	"
V_{11}	$.156 \times 10^{-3}$	"
Fluid Flow Constants		
AF_5	.004	"
F	$.126 \times 10^{6}$	kg/litre/day

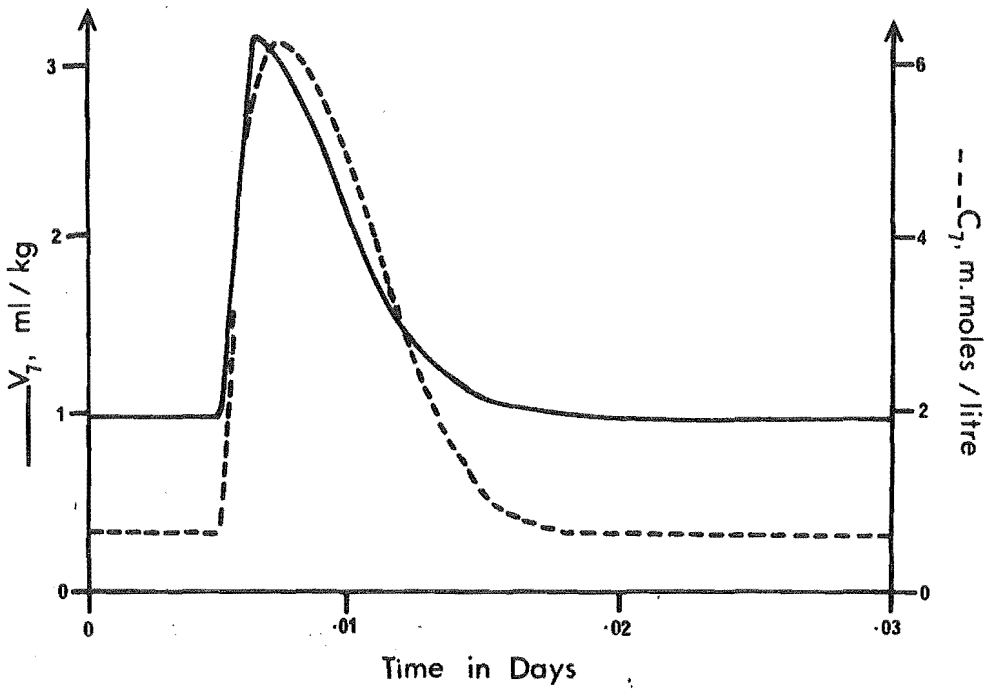


Fig. 9.5 The volume and calcium concentration of compartment 7 predicted by CAMET3 (cf. Fig. 9.1, p.171).

Table 9.6

Results from the Simulation of a Three Meal Day

Quantity	CAMET2	CAMET3	Experimental Value
Initial fasting flow of fluid from the distal ileum (ml/minute)	.36	.36	less than 1 (Whalen et al., 1966)
Accumulated faecal fluid (litres)	.722	.67	approximately .5 (Davenport, 1966, p.171)
Initial fasting concentration of calcium in the flow distal from the ileum (m.moles/litre)	.25	.26	1.5 (Quoted by Pearson, 1972)
Accumulated faecal calcium (mg)	380	408	448 (calculated from the regression line of faecal calcium versus calcium intake given by Malm, 1958)

The predicted values for calcium excretion in the faeces given in Table 9.6 are lower than those occurring in nature. This is due to an elevated calcium absorption which, Pearson (1972) reports, is characteristic of a number of simulations with CAMET2. Although in CAMET3 there is more faecal calcium than in CAMET2 (and hence less calcium is absorbed), the agreement with experiment remains unsatisfactory. The model (9.23) describing the intestinal absorption of calcium was derived by Livesey (1970) from data provided by the perfusion studies of Wensel et al. (1969). These equations and parameters should be reassessed in the light of all the available experimental evidence, including the information about the physical state of calcium in the intestine, currently being studied. One of the reasons Livesey (1970) incorporated the inconsistent fluid equations (9.17)-(9.20) (p.173) into CAMET was to reduce the computer time required for its simulation. The CAMET3 model (Table 9.3, p.174) requires the numerical solution of six more differential equations. However, the time to simulate a three meal day with the entire CAMET3 model (i.e. Table 2.3 (p.25) with the modifications in Table 9.3) using the computer program described in Chapter 10, is only 14 seconds more than the 68 seconds required by CAMET2.

9.2 THE DELAYED APPEARANCE IN BLOOD OF ORALLY GIVEN CALCIUM TRACER

A delay in the appearance in blood of measurable quantities of ingested radioactive calcium has been noted in man by Birge et al. (1969) and Caniggia et al. (1963), and in goats by Gibbons et al. (1972). Typically, 10 to 20 min. pass

after ingestion, before the tracer in the blood rises to .2% of the dose per litre. Caniggia et al. (1963) measured a delay of 20-30 minutes in 13 osteoporotic patients and 10-20 minutes in 5 normal subjects, but other authors have not found such a distinction. Further tests are required to assess the potential of the delay time as a diagnostic index.

The delay undoubtedly occurs, at least in part, during the transport of calcium from the small intestine to the blood. Birge et al. (1969) measured a 10 min. delay on a patient after inserting the tracer directly into the duodenum via a tube, and Gibbons et al. (1972) measured a similar delay after inserting tracer directly into loops of a goat's small intestine. The cause of the delay they measured is uncertain, but there is either an accumulation of calcium in the tissue of the intestinal wall, or a controlled barrier to the passage of calcium. Sampling of both intestinal contents and blood may isolate the correct explanation.

In other experiments Birge et al. (1969), like Caniggia et al. (1963) allowed their subjects to ingest the calcium tracer normally, so that, added to the delay in movement of tracer across the intestinal wall, is the time for tracer to travel from the stomach, where negligible absorption of calcium occurs (Birge et al., 1969), to the small intestine.

In CAMET2, the rate at which material leaves the stomach is given by equation (9.12) (p.173) as the square root of the volume of the stomach contents. This relationship was fitted to the data of Hunt and MacDonald (1954) by Hopkins (1966), who found that it represents their results better than if the emptying rate is proportional to the volume of the stomach contents. The square root law for stomach emptying has little

physical basis. In fact, the rate at which the stomach empties is controlled by receptors, mainly external to the stomach, whose actions are mostly inhibitory (cf. Hunt, 1959). Predominant among the receptors are those in the duodenum, which respond to properties such as the volume, acidity and osmotic pressure of the gastric efflux. The purpose of this inhibitory control may be to protect the duodenal mucosa from the acid gastric efflux, or to regulate the availability of substances (especially carbohydrates; cf. Hunt and Knox, 1967) for absorption. There is no evidence to suggest that duodenal inhibition of gastric emptying acts specifically to control calcium uptake.

To reduce empiricism in CAMET3, we replace the square root law in CAMET2 (equation (9.12)) by equation (9.27) (p.174), which describes the inhibition of stomach emptying as the volume of the upper duodenum increases. We see in Figure 9.7 that there is little quantitative difference between the behaviour of these two models, especially in view of the wide

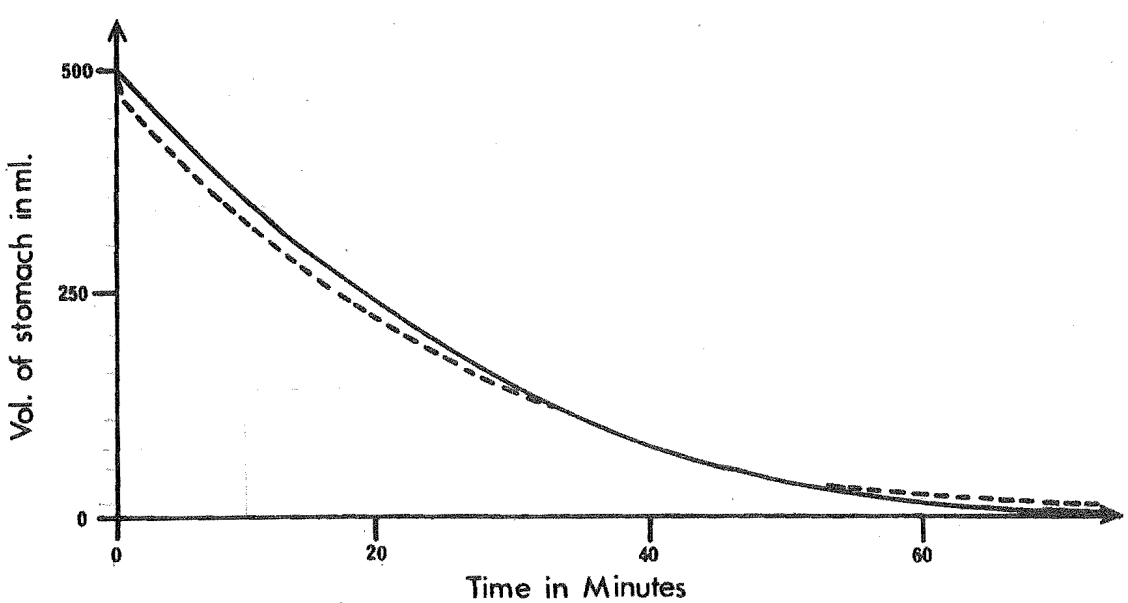


Fig. 9.7 A comparison of two models for stomach emptying rate. —— square root law (9.12); --- duodenal inhibition (9.27).

variation in emptying rate which occurs in nature (cf. Hopkins, 1966). The description of stomach emptying rate (9.27) and the parameter values in Table 9.4 apply only for liquid meals . A more realistic model would allow the parameter values to change according to the properties of the meal as they are known to affect stomach emptying (cf. Hunt and Knox, 1967).

The condition of the stomach at ingestion may affect the time taken for material (such as calcium tracer) to arrive at absorption sites in the intestine. In their serial test meal studies, Hunt and MacDonald (1954) flushed out the stomachs of their patients with water, before the test dose was given. Thus the stomach emptying mechanism may have been stimulated prior to the test, although recently Hunt (1973) reports that there is no noticeable delay in the onset of emptying of test doses of 200 ml of water, given without first washing out the stomach. The tracer doses given by Caniggia et al. (1963) were dissolved in only 10 ml of a calcium gluconate solution, and those given by Birge et al. (1969) were dissolved in 50 ml of skim milk. We know of no experiments which demonstrate the way the stomach empties after such small doses, which are similar in size to the fasting stomach volume (cf. Glass, 1968, p.46).

To assess the effect of stomach emptying on the appearance of tracer in blood, we simulate the experiment of Caniggia et al. (1963) under two conditions. Firstly, we allow the stomach emptying rate $RF_{5,6}$ to vary with volume as in equation (9.27) (p.174) and secondly we hold $RF_{5,6}$ constant at its level at fasting. The dose given by Caniggia et al. (1963) consisted of tracer and 88 mg of calcium dissolved in 10 ml of water.

Thus we set the initial values $M_5(0)$ = .039 m.moles/kg, $V_5(0)$ = .273 ml/kg and $MT_5(0)$ = 1. (We put MT_5 = 1 for convenience because the model is linear with respect to tracer).

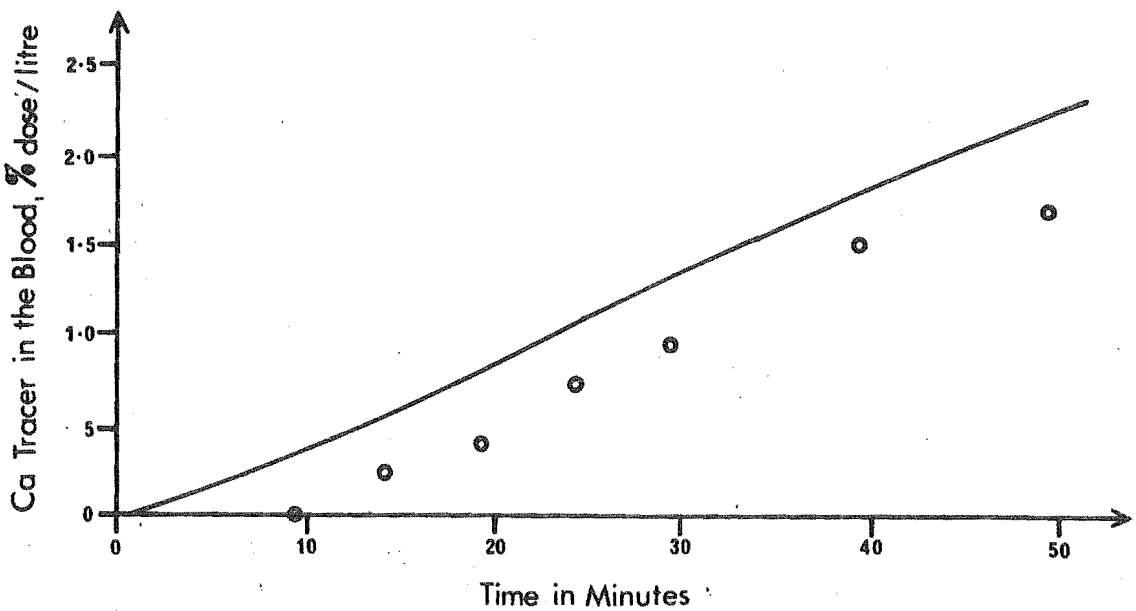


Fig. 9.8 The predicted appearance of orally ingested calcium tracer in blood (——) compared with the averaged data of Caniggia et al. (1963) (OOO).

The results of the two simulations are indistinguishable in Fig. 9.8 and do not account for the delay evident in the averaged data from the five normal subjects tested by Caniggia et al. (1963). We suggest therefore, that the stomach is unlikely to contribute significantly to the delayed appearance of calcium tracer in the blood, unless there is a complete cessation of emptying for a short period following ingestion.

C H A P T E R 1 0

THE SIMULATION OF CAMET3

CAMET3 is implemented on a digital computer, because our analogue computer (EAI 580) cannot accommodate the thirty or more differential equations in it. The program, which is written in FORTRAN for the IBM 360/44, occupies approximately 16,000 words of 32-bit computer memory. It is also acceptable to the Burroughs B6700 computer. Current listings are available from Dr W.S. Metcalf, Chemistry Dept, University of Canterbury, Christchurch, New Zealand.

10.1 PROGRAM ORGANISATION

Each simulation consists of three phases: the option setting phase; the simulation phase; and the graphing phase. During the option setting phase, the user specifies on cards, the parameters of the model, and various parameters to control the operation of the program. Each model parameter (current values are given in Table 2.4 (p.31) and Table 9.4 (p.176)) is identified by a number, as are the control parameters which are associated with the numerical solution of the differential equations (see section 10.2 and Table 10.3, p.191). Other control parameters are set using cards with a code name at the beginning. The various code names, their function, and other information required on the same or subsequent cards are detailed in Table 10.1.

Table 10.1

Control Parameters Identified by a Specific Card Code

Card Code	Function of Card	Information Required
NFLUID NCALC NTRACE NOTHER	To specify the numbers of fluid, calcium, tracer and other differential equations. Other differential equations include those for the hormones PTH and CT, and bone collagen.	The number of equations in each category.
PLOT	To specify the variables to be plotted during the graphing phase.	The number of the graph (1 or 2) and the numbers which identify the variables to be plotted.
WRITE	To specify the variables to be printed out during the simulation phase.	The time interval at which printing is required and the numbers which identify the variables to be printed.
INPUT	To specify the details of an input to the model.	The number which identifies the input, and the details of the size of the doses, the time they start, and their duration.

The parameters NFLUID, NCALC, NTRACE and NOTHER allow the user to exclude differential equations from a particular simulation, resulting in a saving of computer time. For example, if NTRACE is set to zero, the progam does not attempt to solve any of the tracer equations. If NTRACE is set to 11, that number of tracer differential equations are solved.

A card with INPUT at the beginning, is used to specify an input to the model, such as a calcium meal (R_{diet}) or a calcium infusion (R_{inf}). The time the dose is to start, its duration and its magnitude are specified on subsequent cards. Inputs defined for one simulated day, may be repeated for any number of days, after which the regime may be changed if desired. The program sets all model inputs to zero by default.

Cards with WRITE or PLOT at the beginning are used to specify the variables to be printed out during the simulation, or plotted on graphs at the end of it. The values of up to 30 variables at specified time intervals may be printed out, and two graphs, each of up to 8 variables may be plotted on the printer at the end of the simulation. During the simulation, the program saves values of the variables to be plotted. The scaling of the graphs is automatic, but the user can specify whether the origin is to be suppressed or not, whether a logarithmic or linear ordinate is to be used, and whether any of the variables are to be plotted on the same scale, for comparison.

A flow chart of the simulation phase is shown in Figure 10.2. After setting the initial values of the solutions to the model differential equations and after setting the model inputs, the values of all the rates at t = 0 are computed.

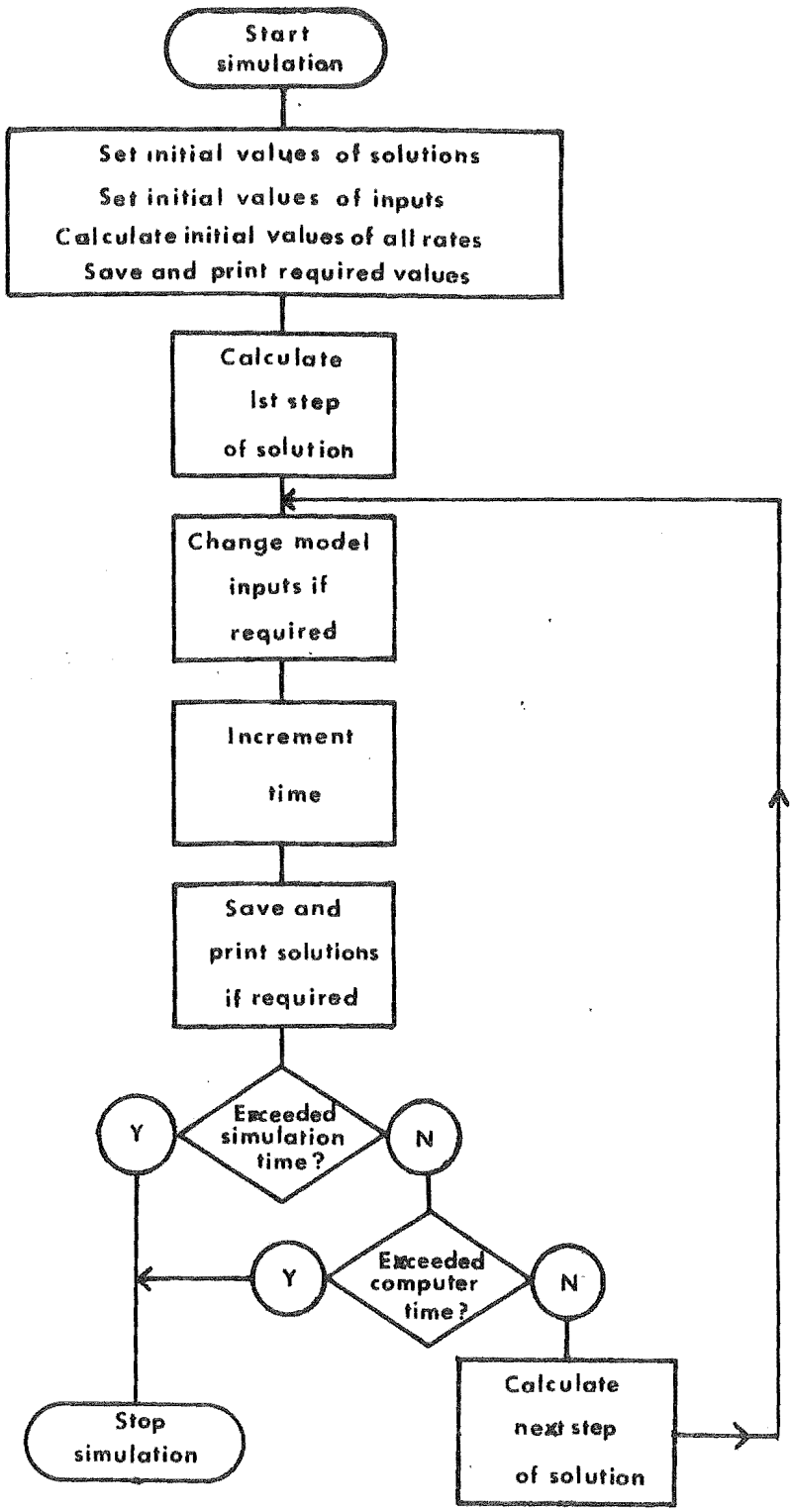


Fig. 10.2 Flow chart of the simulation phase.

The solution then proceeds as a series of small increments in time. The results are printed or saved for graphing, and the inputs are changed as specified during the option setting phase. The simulation terminates when either the independent variable, time, reaches the required value, or a specified amount of computer time has been used. In either case the graphing phase commences immediately.

The equations of CAMET3 are arranged in a subroutine called DIFEQN which is entered whenever the values of the model differentials and rates are computed. To make use of the parameters NFLUID, NCALC, NTRACE and NOTHER, the equations are grouped so that, for example, all the tracer equations are together, and if NTRACE is zero, they can be conveniently omitted from the computations. In the organisation of DIFEQN, care has been taken to ensure that all rates are calculated before they are required to evaluate the differentials. In addition to the differential equations listed in Table 2.3 (p.25) and Table 9.3 (p.174), DIFEQN includes differentials which, from the beginning of the simulation, allow the cumulative totals of a number of model rates to be calculated.

10.2 SOLUTION OF THE DIFFERENTIAL EQUATIONS

The CAMET3 differential equations are solved using a 'predictor-corrector' approximation. At each time step the algorithm predicts a solution from information currently available, and subsequently corrects it after substituting the prediction back into the model differential equations. The differential equations in CAMET3 are all of the form

$$\dot{y} = \mathcal{F}(t,y,---), \tag{10.1}$$

where y is a solution and $\dot{y}$ represents differentiation with respect to time t. The particular predictor-corrector formulae we use to solve (10.1) are a modification of the simple predictor (cf. Kelly, 1967, p.186)

$$y^p_{n+1} = y_{n-1} + 2h\dot{y}_n \tag{10.2}$$

and the modified Euler corrector

$$y^c_{n+1} = y_n + \frac{h}{2}(\dot{y}_n + \dot{y}^p_{n+1}), \tag{10.3}$$

where we denote by y_n the solution at the n^{th} time step. h is the step length. Formulae of the type (10.2) and (10.3) have the advantage that the difference between the predicted and corrected values provides an estimate of the error in the approximation. The step length h may be increased if the error is small, or decreased if the error is unduly large. However, if h is changed, solutions will not be available at the appropriate times to allow (10.2) to be directly applied for the next prediction. Commonly this problem is overcome by interpolating between solutions. In the derivation which follows, we use the integration formula itself to calculate any interpolation.

Let h be the minimum step length, and let y_n by the solution at the n^{th} minimum step (i.e. the solution at t = nh). At time nh, let the previous calculated solution be available at time $(n-n_b)h$, and we propose to calculate the solution at time $(n+n_f)h$. Using the classical Taylor series expansion (cf. Morse and Feshbach, 1954, p.375)

$$y_{n+n_f} = y_n + n_f h\dot{y}_n + \frac{(n_f h)^2}{2!}\ddot{y}_n + \frac{(n_f h)^3}{3!}\dddot{y}_n + \cdots + \tag{10.4}$$

and

$$y_{n-n_b} = y_n - n_b h\dot{y}_n + \frac{(n_b h)^2}{2!}\ddot{y}_n - \frac{(n_b h)^3}{3!}\dddot{y}_n + \cdots + \tag{10.5}$$

Eliminating the term in $\ddot{y}_n$ we have

$$y_{n+n_f} = \left(\frac{n_f}{n_b}\right)^2 y_{n-n_b} + y_n\left(1 - \left(\frac{n_f}{n_b}\right)^2\right) + \frac{n_f}{n_b} h\dot{y}_n(n_f + n_b) + \frac{n_f^2 h^3}{3!}\dddot{y}_n(n_f+n_b), \tag{10.6}$$

where we ignore derivatives of order higher than 3. Equation (10.6) excluding the highest order term, is our predictor formula. We see that if we set $n_b = n_f = 1$, (10.6) reduces to the form (10.2). To obtain the corrector formula, we differentiate (10.4) to give

$$\dot{y}_{n+n_f} = \dot{y}_n + n_f h\ddot{y}_n + \frac{(n_f h)^2}{2!}\dddot{y}_n + \cdots + , \tag{10.7}$$

and eliminate the term in $\dddot{y}_n$ between (10.7) and (10.4), to give

$$y_{n+n_f} = y_n + \frac{n_f h}{2}(\dot{y}_{n+n_f} + \dot{y}_n) - \frac{(n_f h)^3}{12}\ddot{y}_n. \tag{10.8}$$

We denote by $y^p_{n+n_f}$ the predicted solution and $y^c_{n+n_f}$ the corrected solution. If E_p and E_c are the errors in the predictor and corrector formulae respectively, then the accurate solution at any time is given by

$$y^p_{n+n_f} + E_p = y^c_{n+n_f} + E_c. \tag{10.9}$$

Table 10.3

Control Parameters Associated with the Solution of the Differential Equations

Parameter	Function	Current Value
DTMIN	The smallest integration step, and time resolution for printing or graphing solutions, and setting model inputs.	$.5 \times 10^{-4}$
DTMAX	The maximum value of n_f such that $n_f D_{tmin} = D_{tmax}$.	1*
ERRHI	The maximum relative truncation error above which n_f is halved.	$.2 \times 10^{-3}$
ERRLO	The maximum relative truncation error below which n_f is doubled.	$.7 \times 10^{-4}$
TFIN	The value of time at which the simulation is to stop.	-
PROTIM	The maximum allowed computer time.	4 minutes

* This parameter is arbitrarily set because the algorithm has not required the imposition of an upper limit to the step size.

We take $\frac{n_f^2 h^3}{3!}(n_f+n_b)\ddot{y}_n$ as an estimate of E_p and $-\frac{(n_f h)^3}{12}\ddot{y}_n$ as an estimate of E_c. Therefore

$$E_p = -2(n_f+n_b)E_c/n_f \tag{10.10}$$

and from (10.9)

$$E_c = (y^p_{n+n_f} - y^c_{n+n_f}) \,/\, (3 + \frac{2n_f}{n_b}), \tag{10.11}$$

← should be $\frac{2n_b}{n_f}$, as in computer program

which is an estimate of the error in the approximate solution $y^c_{n+n_f}$. The formula for E_c is itself only approximate because E_p and E_c both actually contain terms of higher order than h^3, although we assumed otherwise in our derivation. In the computer program for CAMET3 we use the relative error

$$E = \frac{|E_c|}{y^c_{n+n_f}}. \tag{10.12}$$

It is common to allow for the approximations in E by specifying limits E_{rrhi} and E_{rrlo} such that the step length is decreased if $E > E_{rrhi}$, is increased if $E < E_{rrlo}$, and is left untouched if $E_{rrlo} \leqslant E \leqslant E_{rrhi}$. Accordingly the control parameters ERRHI and ERRLO (Table 10.3) must be set by the user during the option setting phase.

The basic operations in calculating one step of the solution are summarised as follows:

1. Predict using

$$y^p_{n+n_f} = (\frac{n_f}{n_b})^2 y_{n-n_b} + y_n\left[1 - (\frac{n_f}{n_b})^2\right] + \frac{n_f}{n_b} h\dot{y}_n(n_f+n_b); \tag{10.13}$$

2. Calculate the predicted derivative from

$$\dot{y}^p_{n+n_f} = \mathcal{F}(t, y^p_{n+n_f}, \text{---}); \tag{10.14}$$

3. Correct using

$$y^c_{n+n_f} = y_n + \frac{n_f h}{2}(\dot{y}^p_{n+n_f} + \dot{y}_n); \qquad (10.15)$$

4. Estimate the error using

$$E = \frac{|y^p_{n+n_f} - y^c_{n+n_f}|}{y^c_{n+n_f}(3 + \frac{2n_f}{n_b})} \qquad (10.16)$$

We see from (10.13) and (10.15) that step length changes are made simply by altering n_f and n_b as appropriate, but we have the additional advantage that, by suitably controlling n_f and n_b, we can obtain a solution at any time nh. Therefore, the parameter h or DTMIN as it is called in the program (see Table 10.3) specifies a time resolution within which the printing of results, the saving of variables for plotting, and the changing of model inputs will occur.

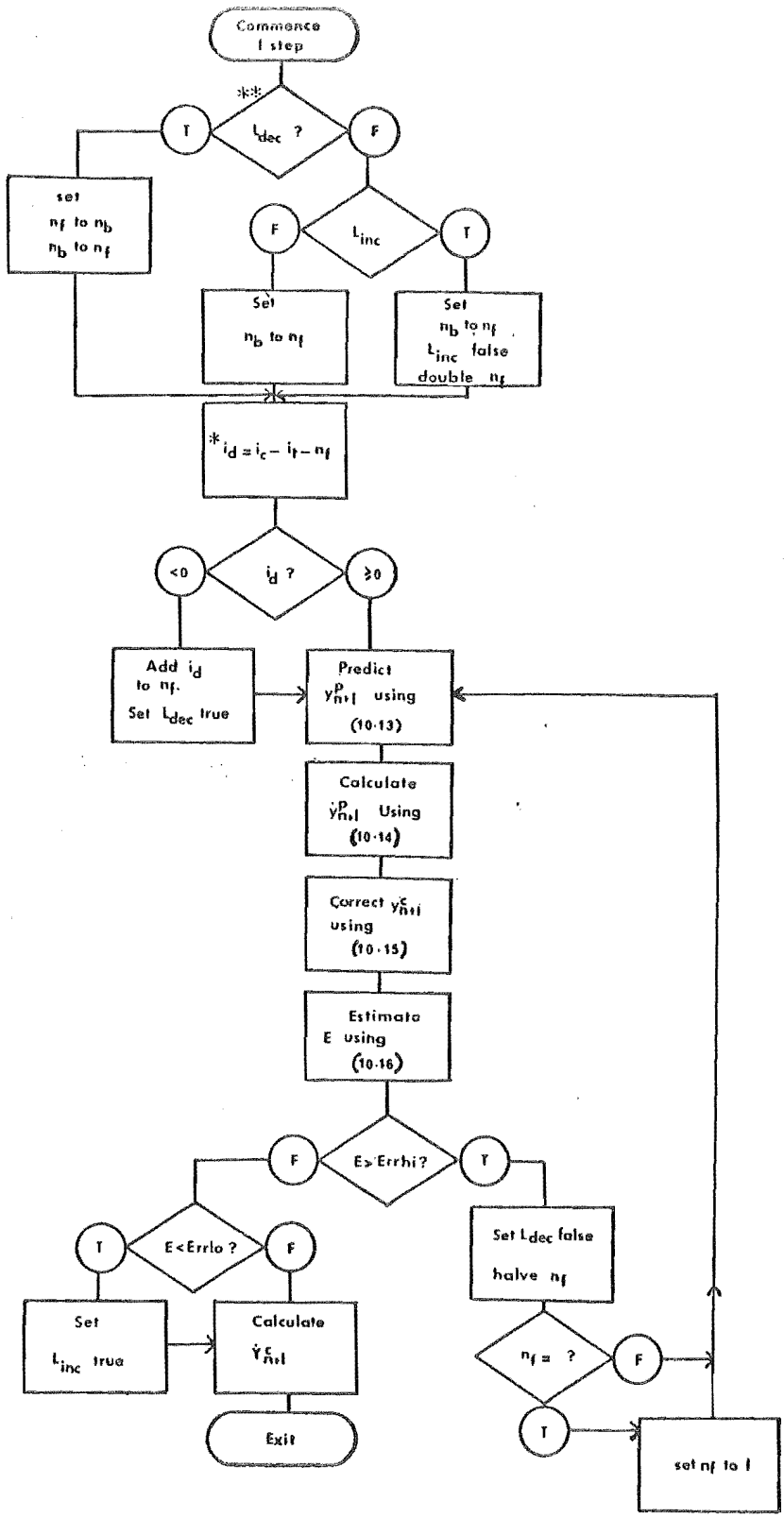
A detailed flowchart of the way these various requirements combine to control n_f and n_b during one step of a simulation is shown in Fig. 10.4. The most important points are:

1. E is calculated for all solutions, and the maximum of these is compared with the error criteria ERRHI and ERRLO.

2. If E > ERRHI, n_f is halved as many times as is necessary to get E $\leqslant$ ERRHI.

3. If E < ERRLO, n_f is doubled for the next step in the solution.

4. If, using the current value of n_f, the next solution will be past a time at which some action is required (eg. printing results), n_f is reduced to give the solution at the



* i_t is time such that $i_t h = t$

i_t is the time t_c of next input save or print such that $i_c h = t_c$

** T represents true, F represents false.

Fig. 10.4 Flow chart showing the step length control logic for one integration step.

desired time. For the next step n_f is set to its value before the interruption, unless 1 and 2 above require it to be reduced.

At the beginning of the simulation y_{0-n_b} is not available (we could not let $n_b = 0$ because it is a divisor in (10.13), and so an alternative method must be used to calculate the first step of the solution. We use the 4th order Runge-Kutta formulae (cf. Kelly, 1967, p.187)

$$\begin{aligned} k_0 &= h\dot{y}(t_0, y_0) \\ k_1 &= h\dot{y}(t_0 + \tfrac{h}{2},\ y_0 + \tfrac{k_0}{2}) \\ k_2 &= h\dot{y}(t_0 + \tfrac{h}{2},\ y_0 + \tfrac{k_1}{2}) \\ k_3 &= h\dot{y}(t_0 + h,\ y_0 + k_2) \end{aligned} \qquad (10.17)$$

and

$$y_1 = y_0 + \tfrac{1}{6}(k_0 + 2k_1 + 2k_2 + k_3).$$

The simulation is started with a step of the minimum length h (i.e. with n_f and $n_b = 1$).

10.3 COMMENTS ON THE INTEGRATION PROCEDURE

The present program is the third which has been written for simulating the calcium metabolism model. The first, called CAMET, (Livesey, 1970) used the approximate formulae (10.2) and (10.3) (p.189) for solving the model differential equations, but was otherwise difficult to control. The second program was written by the present author and was used for the simulation of CAMET2 by Pearson (1972). The integration algorithm was essentially the same as used by Livesey (1970), but the program was structured much as described in section

10.1. However, when simulating with CAMET2, Pearson (1972) found it necessary to adjust ERRHI and ERRLO so that the simulation was carried out using very small step lengths. This was due to a need to ensure that the meal inputs to the model started and stopped at the correct times. The present program guarantees this. To complete simulation 15 which is an intravenous tracer dose followed for three days of three meals each, Pearson (1972) required 16,000 integration steps. The present program, using the CAMET2 differential equations, required only 3000 steps for the same test, while maintaining an accuracy to within three significant figures of Pearson's results.

CAMET3 includes quantities which vary slowly, such as the concentration of calcium in blood, and quantities which vary quickly, such as the stomach volume after the ingestion of a meal. The model is therefore 'stiff' (cf. Gelinas, 1972). The step length in the numerical solution of a stiff model, is controlled by the quantities which vary quickly, even if they have decayed away to very small values. Using more accurate algorithms for solving the differential equations does not always reduce the problem of stiffness. A method may be theoretically more accurate in terms of its Taylor series expansion, but when applied to a stiff model, large spurious solutions or 'instabilities' may be generated. Such was our experience with the algorithm due to Nordsieck (1962). This method is a predictor-corrector procedure which uses the solution and its higher order derivatives at the beginning of a step for the next prediction. Step length changes are very easily carried out, on the basis of tests of the accuracy and stability of the solution. When applied to CAMET2, Nordsieck's

most accurate scheme in which the first Taylor series term omitted is of order 8, gave unstable solutions for some of the rapidly changing variables in the gastro-intestinal section of the model. It is evident that the test for stability in Nordsieck's method is unreliable, a finding supported by the theoretical considerations of Lewis and Stovall (1967).

The present method does not become unstable (probably because it is of low order) and is sufficiently accurate for our purposes. However, it has not been shown to be the best algorithm available in terms of the time required to compute a solution of stated accuracy. The current literature does not give a clear picture as to the particular solution methods which are best suited to a particular model. Current practice is to assess methods by trial. Hull (1969) suggests that there is a need for clearer definition of the problems which occur in solving differential equations, and better assessment of the abilities of methods to handle the problems. As far as the present simulation program is concerned, the 'stiff' methods of Gear (1971) and Treanor (1966) may significantly reduce the computer time required, although their effectiveness may be reduced by the meal inputs which are a part of nearly every simulation.

C H A P T E R 1 1

AN INTERACTIVE DIGITAL SIMULATION PROGRAM

The complete model CAMET3 is large, and the program for its solution described in chapter 10, requires a computer (such as the IBM 360/44 or the Burroughs B6700) which is large enough and fast enough to give answers in an acceptable time. Both the IBM 360/44 and the Burroughs B6700 at the University of Canterbury are currently run in 'batch' mode, whereby the results of a computation are received some hours after the submission of program cards and data. Batch mode does not allow a modeller to interact efficiently with his model. Analogue computers allow such interaction, and solve the differential equations many times faster than do digital computers. However, the scaling of model equations is tedious, as is the wiring up of the analogue components. Further, analogue computers do not provide automatic documentation of the model and tests conducted on it. These objections spoil any advantages of the analogue computer for simulations similar to the single runs reported in chapter 9.

The EAI 640 computer in the Electrical Engineering Dept, University of Canterbury, has 16,000 words of memory, a high speed disc, a magnetic tape drive, a teletype, and a graphic display with a hard copy unit, all of which make it suitable for interactive work. As both the present research and that of Jordan (1973) would directly benefit, a joint project to develop a simulation package for the EAI 640 was undertaken.

Brennan (1967) reviews the programs available for the digital simulation of continuous models. CSMP and DSL have

been used extensively. Interactive packages such as BIOMOD (Groner, 1971), ISL-8 (Benham R.D., 1971), and SIMCON (Anderson et al., 1970) have appeared more recently. Although the present program, SIMUL8, has features in common with some of these, it is designed specifically to fulfil the needs of the present research and that of Jordan (1973).

SIMUL8 is mostly written in FORTRAN IV (EAI, 1971). Assembler language (EAI, 1972) is used where significant savings in either computer memory or computation time can be obtained. An instruction manual and program listings are available from W.K. Kennedy, manager of the hybrid computer laboratory, Electrical Engineering Dept, University of Canterbury, Christchurch, New Zealand.

SIMUL8 consists of a number of utility programs (Table 11.1), which perform various operations such as displaying and setting model parameters, solving the model differential equations, and graphing solutions on the display. Each utility is activated by a command from the teletype as indicated in Figure 11.2.

11.1 SPECIFYING AND SOLVING THE MODEL

We note in Fig. 11.2 that the utility for solving the model equations calls a FORTRAN subroutine DIFEQN which the user must supply to calculate the derivatives, with respect to time, of the dependent variables in his model. Numerical values are trasnferred to and from DIFEQN in three COMMON areas; EQTNS for the model dependent variables, DERIV for their derivatives, and MODPAR for the model parameters.

Table 11.1

SIMUL8 Utilities

Teletype Command	Function
	Simulation Control Parameters*
NS	Set the maximum number of integration steps.
NE	Set the number of model differential equations.
DT	Set the integration step length.
TF	Set the time at which the simulation is to stop.
SS	Solve the model equations.
	Initial Conditions and Model Parameters*
YØ	Set an initial condition.
YY	Investigate the current value of a solution.
MP	Set a model parameter.
RY	Read initial conditions from paper tape.
RP	Read model parameters from paper tape.
OY	Punch initial conditions on to paper tape.
OP	Punch model parameters on to paper tape.
	Plotting
PR	Specify the variables to be saved, during the simulation, for subsequent plotting.
DA	Erase the screen, draw a grid and plot the variables with independent scaling.
DF	As for DA except that all variables are scaled the same as the first variable.
DO	Plot variables on the current graph with the scale factors last used.
DN	Plot variables on the current graph with independent scaling.

Dumping and Reading Data

ED	Read data from paper tape into an array suitable for plotting.
LD	List a saved solution on the teletype, display, or paper tape punch.

Miscellaneous

CT	Type comments on the teletype.
CD	Type comments on the display.
ME	Calculate the root mean squared error between data read in using the ED command, and a model solution.

---

* Where parameters are being changed from the teletype, SIMUL8 displays the present value before accepting the modification.

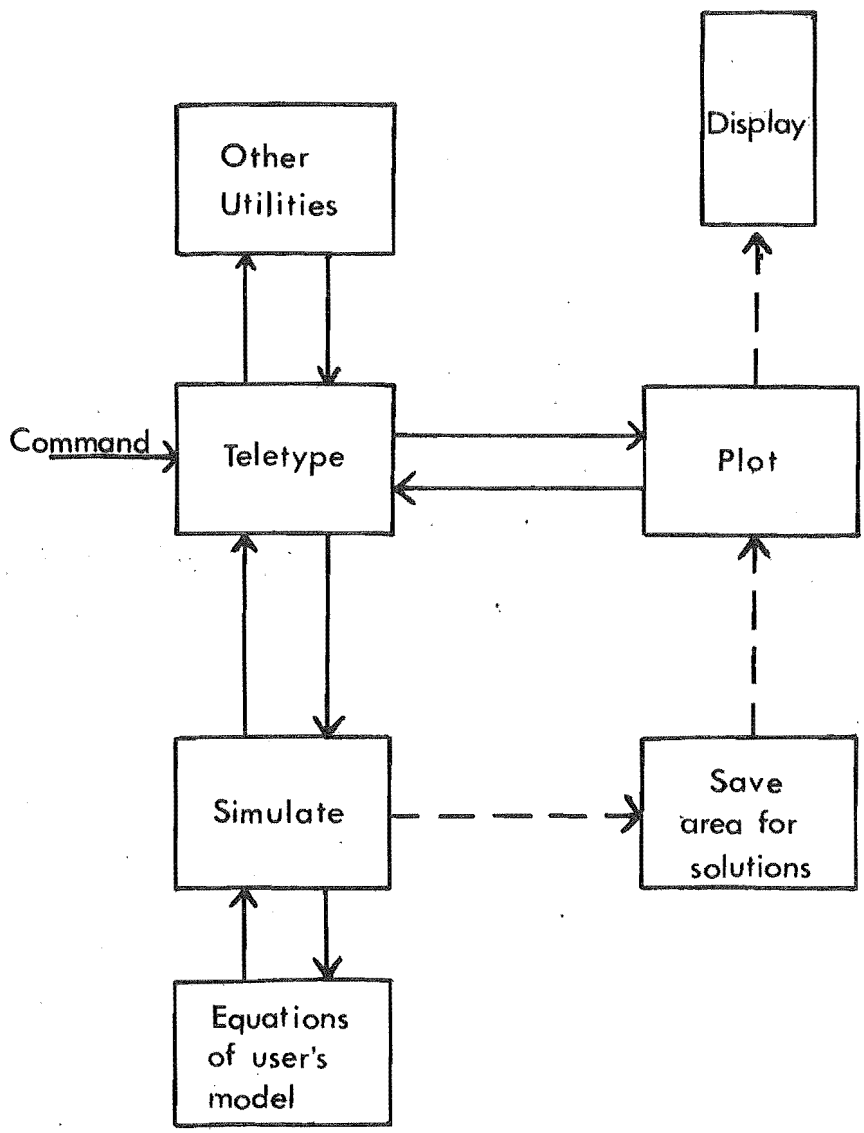


Fig. 11.2 The structure of SIMUL8.

Table 11.3 shows how a model consisting of two differential equations is written as a DIFEQN subroutine for solution with SIMUL8. The straightforward translation from model to subroutine is apparent. The variables and parameters are referenced by the order in which they appear in their respective COMMON areas, so that the quantity B in Table 11.3, is referred to by SIMUL8 as model parameter number 2.

Table 11.3

An Example DIFEQN Subroutine

Model Equations	DIFEQN Subroutine
	SUBROUTINE DIFEQN
	COMMON/EQTNS/Y1,Y2,R1,R2
	COMMON/DERIV/DY1,DY2
	COMMON/MODPAR/A,B,C,D
$r_1 = -by_2$	R1 = -B * Y2
$\frac{dy_1}{dt} = ay_1 + r_1$	DY1 = A * Y1 + R1
$r_2 = dy_1$	R2 = D * Y1
$\frac{dy_2}{dt} = r_2 + cy_2$	DY2 = R2 + C * Y2
	RETURN
	END

The source version of DIFEQN shown in Table 11.3 must be separately compiled before loading it into memory with the other SIMUL8 utilities. So that the minimum of user action is required to edit (on the display), compile and load his DIFEQN subroutine, SIMUL8 automatically controls the various programs

which carry out these functions. It is a simple task to alter DIFEQN, and so change the equations of a model.

11.1.1 Errors in DIFEQN

An important logical rule in writing DIFEQN is that all quantities used in a particular calculation must have been previously defined. For the example in Table 11.3, it is necessary that R1 be calculated before DY1 (which depends on R1) and R2 be calculated before DY2 (which depends on R2). Failure to correctly order the equations results in inaccurate solutions. Simulation languages such as CSMP (Brennan, 1967) carry out a 'presort' of the user's equations to arrange them in their correct order, but we only give a warning if incorrect order is detected. The procedure we use relies on the following facts.

1. The FORTRAN compiler does not set variables to any value when allocating storage. Hence, when the program is loaded into memory, the cells to be used for storing the problem variables are not altered.
2. To carry out an arithmetic operation such as add or multiply on the EAI 640, a link is set up from the user's program to a standard arithmetic subroutine. That is, the arithmetic operations are 'out of line'.

Before loading the SIMUL8 utilities and the user's DIFEQN we set every computer memory location to 255, one of many 'invalid' numbers in the floating point representation used in the EAI 640 (EAI, 1971). Therefore, if the number 255 is encountered during an arithmetic operation, we know that a variable has not been previously defined, thus violating our requirements for equation order. An error message results. In addition,

the above procedure indicates if a model parameter or initial condition has not been set.

11.1.2 Solution of the Differential Equations

Unlike the simulation program for CAMET3 (chapter 10, p.192), SIMUL8 uses a fixed step length algorithm for the solution of the model differential equations. The algorithm, originally published by Bashforth and Adams (1883), is given by (cf. Henrici, 1962)

$$y_{n+1} = y_n + h(\tfrac{3}{2}\dot{y}_n - \tfrac{1}{2}\dot{y}_{n-1}), \qquad (11.2)$$

where y_n is the solution y at time nh, h is the step length, and

$$\dot{y}_n = \mathcal{F}(t, y_n, \text{---}). \qquad (11.3)$$

This algorithm has so far proved satisfactory. It is recommended that before a solution is accepted, it should be repeated with a reduced step length.

11.2 PLOTTING

When output is displayed on a storage oscilloscope, it remains until the whole screen is 'erased'. Consequently, the graphical solutions from a number of simulations may be superimposed for comparison. The basic graphing program in SIMUL8 takes an array containing samples of the solution for a particular model variable, scales it, suppresses the origin (optional) and graphs it on the screen. Unfortunately, the solution may not be graphed as it is calculated because scaling information cannot be determined automatically, until all the values to be plotted are available. Therefore, during

the course of a simulation SIMUL8 stores samples of any four of the variables (previously chosen by the user) in EQTNS COMMON. To allow full flexibility in comparing solutions, we use the following four plotting commands (N1, N2, N3 and N4 refer by number to variables in EQTNS COMMON which have been saved for plotting during the previous simulation.)

DA ← N1,N2,N3,N4 Erase the display screen, draw new graph axes, and plot each of the variables listed with independent scaling and origin suppression.

DF ← N1,N2,N3,N4 Erase the display screen, draw new graph axes, and plot each of the variables listed but using for N2, N3 and N4 the same origin and scale factor as determined automatically for variable N1.

DO ← N1,N2,N3,N4 Plot the variables indicated on the current axes using the same origin and scale factor as for the previously drawn graph.

DN ← N1,N2,N3,N4 Plot the variables N1,N2,N3,N4 on the current axes but using independently determined origins and scales for each variable.

It is necessary to specify only those variables N1 to N4 which are to be plotted.

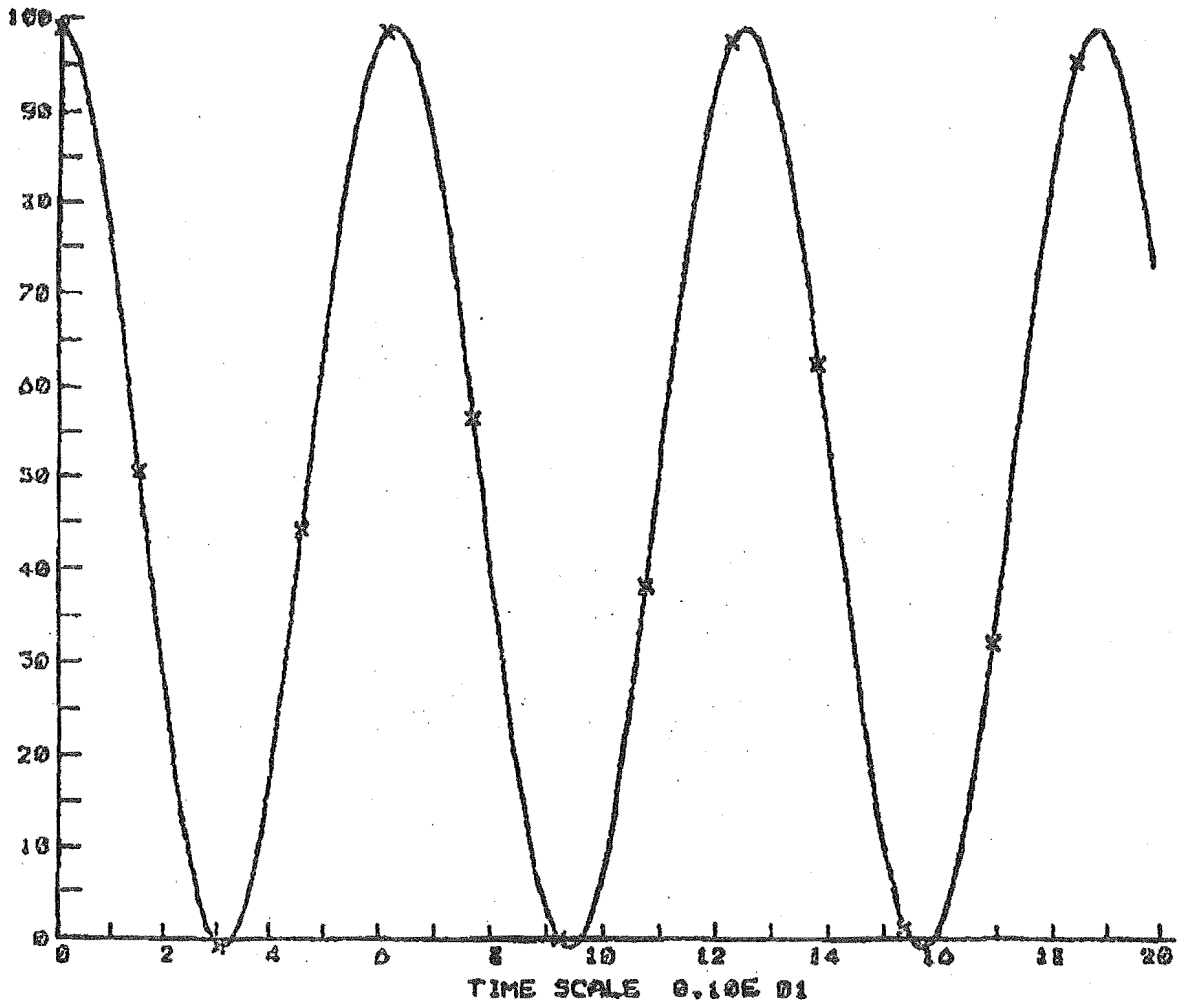
11.3 CONCLUSIONS

To demonstrate the use of SIMUL8, we solve the simple model shown in Table 11.3 (p.203) with $y_1(0) = 1$, $a = c = y_2(0) = 0$. The analytical solution for $y_1(t)$ is

$$y_1(t) = \cos \sqrt{bd}\, t \qquad (11.4)$$

A copy of the teletype printout obtained during this simulation, is given in Appendix II. After some manipulation of the parameters b and c the display copied in Fig. 11.4 was obtained. The theoretical solution (11.4) which we generated simultaneously, is shown for comparison.

SIMUL8 is machine particular in detail only, although the scheme described in section 11.1.1 for checking errors in the model may not be possible, in its present form, on a different computer. The EAI 640 does not have fast floating point arithmetic; its time for a multiplication is approximately 20 times that of the IBM 360/44. Therefore we cannot expect to use SIMUL8 for large models such as CAMET3. The efficiency of simulation depends both on the speed of the computer and on the time required to prepare the model for testing. This is especially so for the biological models of the sort discussed in chapter 9 and in Jordan (1973). The straightforward correspondence between model equations and DIFEQN subroutine evident in Table 11.3 has made it easy for a number of users of SIMUL8 to obtain solutions to their models conveniently.



SYMBOL	VARIABLE	RUN NO.	SCALE	ORIGIN	MAXIMUM	MINIMUM
———	1	0	0.20E-01	0.49E 02	0.100E 01	-.997E 00
X——X	3	0	0.20E-01	0.49E 02	0.100E 01	-.100E 01

GRAPH 3 SIMUL8 DEMONSTRATION 25/9/73

Fig. 11.4 Graphical output from SIMUL8.

CHAPTER 12

CONCLUSIONS

12.1 EXPONENTIAL FITTING

As long as compartment models are used to represent the results of tracer studies, the problem of fitting decaying exponential functions to data will remain. In chapters 3 and 4 we demonstrate that an improvement in the estimates of the number of compartments and the exponential coefficients and exponents can only be obtained if the accuracy of the data is improved. We must accept that limitations in the estimated parameters reflect, in the main, limitations on the amount of information in the data, and not a deficiency in the fitting procedure.

12.2 INVERSE FILTERING IN THE COMPLEX PLANE

The results obtained in section 5.1 indicate that this procedure may improve some deconvolution problems. This contention could be further assessed by a theoretical treatment of the behaviour of the Fourier transform of data error in the complex plane. Tests on the method for different functions $g(x)$ and $h(x)$ would provide useful information. Using present fast Fourier transform methods, it would be quite feasible to choose the best contour in the complex plane by trial.

12.3 CONSISTENT DECONVOLUTION

The consistency conditions for the deconvolution of functions, which were introduced in section 5.2.1 provide, in principle, a powerful method for assessing and improving the accuracy of the data.

However, in chapter 6 we successfully applied this theory only to ideal data. The lack of convergence of the positions of the zeros calculated from the experimental data in section 6.4, suggests that Hermite functions may be unsuitable for representing this particular data. Apart from increasing the efficiency of the computations, it is unlikely that the alternative method described in section 8.3.2 for fitting the Hermite functions to the data, would improve the deconvolution procedure. Better results may be obtained if the experimental data is assumed to be of finite extent, and if the zeros are calculated by first representing it by a Fourier series.

In chapter 7 we successfully applied the consistency conditions to the deconvolution of both analytically generated and actual experimental data. This method of consistent data completion could be applied to other problems of the same type; such as the removal of catheter 'smearing' of blood samples (cf. Branston, 1972), and the removal of the effects of recirculating indicator which often limits the amount of information which can be derived from tracer studies of the circulation (cf. Coulam et al., 1967). A disadvantage of Fourier deconvolution methods is that it is difficult to impose constraints on the result. In section 7.4 we have instances where there is a flow of calcium tracer from gut to blood even before the tracer has been introduced into the

system. This unreal situation is a consequence of the filtering of the Fourier transform. The concept of 'entropy' introduced by Frieden and Burke (1972) goes some way towards imposing a positivity constraint on deconvolved functions, but more work is required in this area.

The deblurring of pictures which have been distorted by linear motion, is a deconvolution problem in which the zeros of G(w) all lie along the real line (cf. Bates et al., 1973). If the complete blurred picture is not available, the deconvolution techniques of chapter 7 apply. But because the zeros of G(w) lie along the real line, the matrix inversion used in section 7.2 to complete the data may profitably be replaced by the following iterative scheme.

The Fourier transform of the truncated data is calculated and forced to zero at the known zeros of G(w). f(x) is reconstructed from its modified transform. In the area where f(x) is given originally it will now not agree with the data. It should be replaced by the data and the process repeated until the f(x) reconstructed from its modified transform is within a specified accuracy of the data. The completed f(x) can then be deconvolved. Current deblurring methods are not successful when the blurred picture is incomplete (cf. Bates et al., 1973).

12.4 THE MODELLING AND SIMULATION OF CALCIUM METABOLISM

In chapter 9 we removed the logical inconsistency from the gut fluid equations of CAMET2, and included a more physical description of the emptying of the stomach. A further inconsistency in the model is due to the constant plasma volume assumed. When a meal is ingested the model

predicts a drop in the amount of calcium in the blood caused by the raised intestinal secretions. Whereas in nature the volume of the blood compartment also falls for the same reason, the model maintains a constant plasma volume, resulting in a transient drop in blood calcium concentration. Further inadequacies of the model as noted by Pearson (1972) are mentioned in section 2.2 (p.30).

The observed delay in appearance in blood of orally given calcium tracer is unexplained. This problem deserves some attention. Firstly because of the suggestion by Caniggia et al. (1963) that the delay time is longer for osteoporotic patients, and secondly because the mechanism of the delay may provide clues as to the control of calcium absorption.

The simulation program for CAMET3 is an advance over that for CAMET and CAMET2. It is faster, and more easily used. Even so, a simulation of, for instance, three weeks, may take an unacceptably large amount of computer time. The efficiency of algorithms especially designed for solving 'stiff' systems of differential equations should be tested in this respect. As far as the structure of the program is concerned, it seems likely that remote, interactive, computer terminals will soon be available at this university and so consideration should be given to modifying the program to take advantage of these facilities. The design principles of SIMUL8 might well be applied.

APPENDIX I

LIST OF SYMBOLS USED IN THE CAMET3 MODEL

The following upper case symbols are used in the model equations to denote concentrations, rates etc.

C	concentration
F	fluid flow variable
M	quantity
R	rate
S	cumulative quantity
SA	specific activity
T	tracer calcium

Extensive quantities are normalised to unit body weight.

$\bar{C}_1$	denotes the mean value of the plasma calcium - in this model taken as 2.5 m.moles/litre
$\dot{x}$	represents the first derivative of x with respect to time (dx/dt)

The symbols used throughout the present work to represent the parameters and variables of the model are as follows:

A_I	empirical constant
AF_I	the fraction of the fluid entering the intestinal compartment I, from compartment (I-1), that is absorbed; except that AF_5 is the maximum volume of compartment 6.
$B_{I,1}$ $B_{I,2}$	parameters controlling the volume of the intestinal compartment I
C_I	the concentration of calcium in compartment I
C_{bal}	the calcium balance

C_c	concentration of calcitonin
C_p	concentration of parathyroid hormone
D_{tmin}	the minimum allowable step length
$E_{I,1}$ $E_{I,2}$	parameters controlling the absorption of calcium from intestinal compartment I
E_{rrhi}	the value of the truncation error above which the step length is decreased
E_{rrlo}	the value of the truncation error below which the step length is increased
F	parameter controlling the rate of stomach emptying
G_I	duodenal secretion parameter
$K_{I,J}$	first order rate constant for the transfer of calcium from compartment I to compartment J
K_{acc} K_d K_{dc} K_{dp} K_{oss}	rate constants
K_c	$C_1/\bar{C}_1$
K_{pr} K_{rp}	proportionality constants between the bone collagen and the bone calcium
M_I	quantity of calcium in compartment I
M_{bca}	quantity of bone calcium
M_{bcoll}	quantity of bone collagen
M_c	quantity of calcitonin
M_p	quantity of parathyroid hormone
MT_{tot}	quantity of tracer calcium in model compartments at t = 0
$R_{I,J}$	rate of unidirectional flow of calcium from compartment I to compartment J

R_{abs}	rate of calcium absorption from the intestinal compartments
R_{acc}	rate of calcium deposition in the bone
R_{cinf}	rate of calcitonin infusion
R_{coacc}	rate of collagen synthesis
R_{codec}	rate of collagen destruction
R_{dc}	rate of calcitonin catabolism
R_{dec}	rate of calcium resorption from the bone
R_{derm}	rate of loss of calcium in the sweat
R_{diet}	rate of ingestion of dietary calcium
R_{dp}	rate of parathyroid hormone catabolism
R_{fecl}	rate of excretion of calcium in the faeces (strictly the rate of flow of calcium into the colon)
R_{gsec}	rate of calcium excretion into the intestinal lumen
R_{inf}	rate of intravenous calcium infusion
R_{pinf}	rate of infusion of parathyroid hormone
R_{sc}	rate of secretion of calcitonin
R_{sp}	rate of secretion of parathyroid hormone
R_{urin}	rate of loss of calcium in the urine
RG_I	basal secretion rate of calcium into the intestinal compartment I
RT_{acc}	rate of tracer uptake by the bone (including incorporation in new bone as well as long term exchange)
S_{acc}	cumulative skeletal uptake of calcium by bone formation
S_{dec}	cumulative resorption of bone calcium
S_{derm}	cumulative loss of calcium in the sweat

S_{diet}	cumulative ingestion of dietary calcium
S_{fecl}	cumulative excretion of faecal calcium
S_{gsec}	cumulative intestinal calcium secretion
S_{inf}	cumulative intravenous calcium infusion
S_{abs}	cumulative absorption of calcium from the intestinal compartments
S_{urin}	cumulative excretion of urinary calcium
SA_I	specific activity of tracer calcium in compartment I
U_{mcoll}	quantity of unmineralized collagen
V_I	volume of compartment I
V_c	volume of distribution of calcitonin
V_p	volume of distribution of parathyroid hormone

APPENDIX II

DEMONSTRATION OF SIMUL8

We use the SIMUL8 program described in chapter 10 to solve the simple model shown in Table 11.3 (p.203) for the conditions $y_1(0) = 1$ and $a = c = y_2(0) = 0$. The FORTRAN which describes the model is shown below, where we note that we have included a calculation of the analytical solution (y_{1act}) given by equation (11.4) (p.206).

```
C
C
C       DIFEQN SUBROUTINE TO DEMONSTRATE
C
C       SIMUL8 BY SOLVING THE MODEL IN
C
C       TABLE 11.3
C
C
C
        SUBROUTINE DIFEQN
C
        COMMON /EQTNS/ Y1,Y2,Y1ACT
C
        COMMON /DERIV/ DY1,DY2
C
        COMMON /MODPAR/ A,B,C,D
C
C       THE MODEL EQUATIONS ARE AS FOLLOWS
C
        R1   =   -B * Y2

        DY1  =   A * Y1 + R1

        R2   =   D * Y1

        DY2  =   C * Y2 + R2
C
C       ANALYTICAL SOLUTION
C
        T    =   TIME (T)
        Y1ACT   =   COS(SQRT(B*D)*T)
C
        RETURN
        END.
```

The following is a copy of the teletype printout obtained during this simulation. The various commands used are summarised in Table 11.1 (p.200), and where necessary notes are added for clarification.

```
   PAGE 01

   E.A.I. 640 DIGITAL SIMULATION

   IDENTIFY← SIMUL8 DEMONSTRATION 25/9/73
   LP
   LI
1) IP
   LI
   IP

   NO. OF STEPS← 1000
    STEP LENGTH← .005
   FINAL TIME← 1.
   VARIABLES TO BE PREPARED ← 1,2,3
   NO. OF DIFF. EQTNS. ← 2
2) INITIAL COND. DEVICE ← 2
    1 ← 1.
    2 ← 0.
   MODEL PARAMETER DEVICE ← 2
    1 ← 0
    2 ← 1.
    3 ← 5.
    4 ← 5.
3)  5 ← X
   X←
   MP ← 3:  0.5000E 01   ← 0
      ← X
   X←
    SS
    RUN NO.  1
   X←
   DA ← 1,3
   X←
4) TF   0.1000E 01   ← 10.
   X←
    SS
    RUN NO.  2
5)      TIME=  0.500E 01
     NO. OF STEPS= 1000
   X←
   DA ← 1,3
   X←
   MP ← 1:   0.0         ←
      ← 2:   0.1000E 01  ←
      ← 4:   0.4000E 01  ← 1
      ← X
   X←
    SS
    RUN NO.  3
        TIME=  0.400E 01
     NO. OF STEPS= 1000
   X←
   DO ← 1
```

```
    PAGE 02

    X←
    TF    0.1000E 02    ← X
    X←
 6) NS     1000    ←  50000
    X←
     SS
     RUN NO.   4
    X←
    DA ← 1,3
    X←
 7) DT    0.4999 E-02    ← .01
    X←
     SS
     RUN NO.   5
    X←
 8) DO ← 1
    X←
    DA ← 1,3
    X←
    TF    0.1000E 02   ←  20
    X←
     SS
      RUN NO.   6
    X←
 9) DA ← 1,3
    X←
```

1. The user's DIFEQN is being loaded.
2. Device number 2 specifies the teletype.
3. The character X causes an immediate return to the command input phase.
4. Increase the final simulated time.
5. The simulation terminates before the final time of 10 because too few integration steps have been allowed.
6. Increase the maximum number of integration steps.
7. Increase the step length.
8. Plot results on top of previous solution for comparison.
9. Plot the graph shown as Fig. 11.4 (p.208).

REFERENCES

Abbrecht P.H. and Prodany N.W. (1971) A Model of the Patient - Artificial Kidney System. IEEE Trans. on Biomedical Eng. BME-18: 257.

Abramowitz M. and Stegun I.A. (1965) Handbook of Mathematical Functions. Dover Publications.

Achinstein P. (1965) Theoretical Models, Brit. J. Phil. Sci. 16: 102.

Ackerman E., Gatewood L.C., Rosevear J.W., Molnar G.D. (1969) Blood Glucose Regulation and Diabetes, in Heinmets F. (ed.) Concepts and Models of Biomathematics, Marcel Dekker.

Ackerman E., Strickland E.H., Hazelrig J.B., Gatewood L.C. (1967) Computers in biomathematical applications, Clinical Pharmacology and Therapeutics 8: 170.

Acton F.S. (1970) Numerical Methods that Work, Harper & Row.

Agarwal G.C., Berman B.M., Stark L. (1970) Studies in Postural Control Systems. Part I. Torque Disturbance Input. IEEE Trans. on Systems Science and Cybernetics SSC-6: 116.

Anderson D.U., Knopp T.J., Bassingthwaighte J.B. (1970) SIMCON - simulation control to optimize man-machine interaction. Simulation 14: 81.

Anderson J.H., Achs M.J., Garfinkel D. (1971) Gluconeogenesis in Rat Liver Cytosol. II. Computer Simulation of Control Properties. Cptrs and Biomed. Res. 4: 107.

Anderson J., Tomlinson R.W.S., Osborn S.B., Wise M.E. (1967) Radiocalcium turnover in man. Lancet i: 930.

Andreae J.H., Cox G.A., Sloane M.R. (1972) Modelling and Simulation. Electrical Engineering Report No. 6. Dept. of Electrical Engineering, University of Canterbury, Christchurch, New Zealand.

Apter J.T. (1970) Biosystems Modelling, in Clynes M. and Milsum J. (eds) Biomedical Engineering Systems, McGraw-Hill.

Arnaud C.D., Tsao H.S., Littledike T. (1971) Radioimmunoassay of human parathyroid hormone in serum. J. Clin. Invest. 50: 21-34.

Åström K.H., Eykhoff P. (1971) System Identification - a Survey. Automatica 7: 123.

Atkins G.L. (1969) Multicompartment Models for Biological Systems, Methuen.

Aubert J.-P., Bronner F. (1965) A symbolic model for the regulation by bone metabolism of the blood calcium level in rats. Biophysical Journal 5: 349.

Aubert J.-P., Milhaud G. (1960) Methode de mèsure des principales voies du métabolisme calcique chez l'homme. Biochimica et Biophysica Acta 39: 122.

Bashforth F., Adams J.J. (1883) Theories of Capillary Action, Cambridge University Press.

Bates R.H.T. (1965) Estimating the Length of a Compact Function from a Noisy Measurement of the Modulus of its Fourier Transform. Electronics Letters 1: 279.

Bates R.H.T. (1969a) Only a Countable Number of Brightness Temperature Distributions Could Produce an Observed Intensity Interferogram. N.Z. J. Sci. 12: 467.

Bates R.H.T. (1969b) Contributions to the Theory of Intensity Interferometry. Mon. Not. R. Astr. Soc. 143: 413.

Bates R.H.T., Napier P.J. (1972) Identification and Removal of Phase Errors in Interferometry. Mon. Not. R. Astr. Soc. 158: 405.

Bates R.H.T., Kennedy W.K., McDonnell M.J. (1973) Efficient digital restoration of images blurred by linear motion. To be submitted to - Letters in Applied and Engineering Sciences.

Bauer G.C.H., Carlsson A., Lindquist B. (1957) Bone salt metabolism in humans studied by means of radiocalcium. Acta Medica Scandinavica 158: 143.

Baylor C.H., Van Alstine H.E., Keutmann E.H., Basett S.H. (1950) The fate of intravenously administered calcium. Effect on urinary calcium and phosphorus, fecal calcium and calcium-phosphorus balance. J. Clin. Invest. 29: 1167.

Beach S.A., Dyson E.D. (1971) The mathematical development and computer simulation of compartment flow models in problems of internal radionuclide metabolism. In Conference on Computers for Analysis and Control in Medical and Biological Research 7-9 Sept. 1971. IEE Conf. Publ. No. 79.

Bekey G.A. (1970) System identification - an introduction and a survey. Simulation 15: 151.

Bellman R., Åström K.J. (1970) On Structural Identifiability Math. Biosci. 7: 329.

Bellman R.E., Kalaba R.E., Lockett J. (1966) Numerical Inversion of the Laplace Transform. American Elsevier.

Benetazzo L., Busnardo B., Clemente G. (1972) A functional model of early thyroxine distribution in man. Med. Biol. Eng. 10: 337.

Benham R.D. (1971) Interactive simulation language - 8. Simulation 16: 116.

Bergland G.D. (1969) A guided tour of the fast Fourier transform. IEEE Spectrum 41: 52.

Berman M. (1963) Formulation and Testing of Models. Annals N.Y. Acad. Sci. no. 108, p.182.

Berman M., Schoenfeld R. (1956) Invariants in Experimental Data on Linear Kinetics and the Formulation of Models. J. Appl. Phys. 27: 1361.

Berman M., Shahn E., Weiss M.E. (1962) The routine fitting of kinetic data to models: A mathematical formalism for digital computers. Biophysical Journal 2: 275.

Berry C.R. (1947) The shape of X-ray diffraction lines from colloidal magnesium oxide. Physical Rev. 72: 942.

Bianchi C.P. (1968) Cell Calcium, Butterworths, London.

Biddulph D.M., Hirsch P.F., Cooper C.W., Munson P.L. (1970) Effect of thyroparathyroidectomy and parathyroid hormone on urinary excretion of calcium and phosphate in the golden hamster. Endocrinology 87: 1346.

Birge S.J., Peck, W.A., Berman M., Whedon G.D. (1969) Study of calcium absorption in man: A kinetic analysis and physiologic model. J. Clin. Invest. 48: 1705.

Bogumil R.J., Ferin M., Rootenberg J., Speroff L., Vande Wiele R.L. (1972) Mathematical studies of the human menstrual cycle. I. Formulation of a mathematical model. J. Clin. Endocrin. and Metab. 35: 126.

Boyers D.G., Cuthbertson J.G., Luetscher J.A. (1972). Simulation of the human cardiovascular system: a model with normal responses to change of posture, blood loss, transfusion and autonomic blockade. Simulation 18: 197.

Branston N.M., Read R.R. (1972) Catheter Smearing Correction Using Inverse Filtering Techniques. IEEE Trans. on

Biomed. Eng. BME-19: 286.

Brennan R.D. (1967) Continuous system modelling programs, in Buxton J.N. (ed.) Simulation Programming Languages North-Holland.

Brennan R.D. (1968) Simulation is Wha-a-at? Part II, in McLeod J. (ed.) Simulation. The Dynamic Modeling of Ideas and Systems with Computers. McGraw-Hill Inc.

Bronner F., Aubert J.-P. (1965) Bone metabolism and regulation of the blood calcium level in rats. Amer. J. Physiol. 209: 887.

Brown K.M. (1972) Computer Oriented Methods for Fitting Tabular Data in the Linear and Nonlinear Least Squared Sense. AFIPS Conf. Proc. 41 Part II: 1309. Fall Joint Computer Conference 1972.

Brownell G.L., Callahan A.B. (1963) Transform Methods for Tracer Data Analysis. Annals N.Y. Acad. Sci. no. 108, p.172.

Burkinshaw L., Marshall D.H., Oxby C.B., Spiers F.W., Nordin B.E.C., Young M.M. (1969). Bone turnover model based on a continuously expanding exchangeable calcium pool. Nature 222: 146.

Burrus C.S., Parks T.W., Watt T.B. (1971) A digital parameter-identification technique applied to biological signals. IEEE Trans. on Biomed. Eng. BME-18: 35.

Campbell G.A., Foster R.M. (1961) Fourier Integrals, D. Van Nostrand N.Y.

Canary J.J., Carreon G.G., Bloomer H.A., Kyle L.H., Meloni C.R. (1962) Serial changes in serum calcium and phosphorus concentration and in urinary phosphorus excretion after parathyroid surgery: Further evidence for a dual effect of parathyroid hormone. J. Clin. Endocrin. and Metab. 22: 229.

Caniggia A., Gennari C., Bianchi V., Guideri R. (1963) Intestinal absorption of ^{45}Ca in senile osteoporosis. Acta Medica Scandinavica 173: 613.

Chen C.F., Shieh L.S. (1968) A novel approach to linear model simplification. Int. J. Contr. 8: 565.

Coddington E.A., Levinson N. (1955) Theory of Ordinary Differential Equations, McGraw-Hill.

Cohn S.H., Bozzo S.R., Jesseph J.E., Constantinides C., Huene D.R., Gusmano E.A. (1965) Formulation and testing of a compartmental model for calcium metabolism in man. Radiat. Res. 26: 319.

Coleman T.G. (1972) Simulation is Helping Biomedical Research. Simulation 19: 29.

Cook G.E., Denman E.D., Carr H.M. (1967) Numerical Inversion of Laplace transforms by the Laguerre-Gauss Quadrature Approximation. IEEE Trans. on Automatic Control AC-12: 623.

Copp D.H. (1960) Parathyroids and Homeostasis of blood calcium. In Rondahl K., Nicholson J.T., Brown E.M. (eds) Bone as a Tissue, McGraw-Hill.

Copp D.H., Mensen E.D., McPherson G.D. (1960) Regulation of blood calcium. Clinical Orthopaedics 17: 288.

Cornell R.G. (1962) A Method for Fitting Linear Combinations of Exponentials. Biometrics 18: 104.

Coulam C.M., Warner H.R., Marshall H.W., Bassingthwaighte J.B. (1967) A steady-state transfer function analysis of portions of the circulatory system using indicator dilution techniques. Cptrg and Biomed. Res. 1:124.

Davenport H.W. (1966) Physiology of the Digestive Tract, Year Book Medical Publishers Inc.

Davis R.H., Ottaway J.H. (1972) Application of Optimization Procedures to Tracer Kinetic Data. Math. Biosci. 13: 265.

Diamessis J.E. (1972) Least-Square-Exponential Approximation Electron. Lett. 8: 454.

Dickinson C.J. (1971) A study of urinary water and sodium excretion using a digital computer model of systemic circulation, body fluid compartments and kidneys. In conference on Computers for Analysis and Control in Medical and Biological Research 7-9 Sept. 1971. IEE Conf. Publ. No. 79.

Dillard R.L., Eastman H., Fordtran J.S. (1965) Volume-flow relationship during the transport of fluid through the human small intestine. Gastroenterology 49: 58.

Dionne P.J. (1972) RECONS --- a REspiratory CONtrol System benchmark simulation. Simulation 19: 73.

Dixon L.C.W. (1972) Nonlinear Optimisation, English Universities Press.

EAI (1971) FORTRAN IV Reference Manual, Electronic Associates Inc., Publ. No. 00827.0022-2.

EAI (1972) Assembler Manual, Electronic Associates Inc., Publ. No. 00827.0015-5.

Elgerd O.I. (1967) Control Systems Theory, McGraw-Hill.

Eykhoff P. (1963) Some Fundamental Aspects of Process-Parameter Estimation. IEEE Trans. on Automatic Control AC-4: 347.

Fahrland D.A. (1970) Combined discrete event continuous simulation. Simulation 14: 61.

Flanagan S.L. (1972) Voices of Men and Machines. J. Acoustical Soc. Amer. 51: 1375.

Frieden B.R., Burke J.J. (1972) Restoring with maximum entropy II: Superresolution of photographs of diffraction-blurred impulses. J. Opt. Soc. Amer. 62: 1202.

Gardner D.G., Gardner J.C., Laugh G., Meinke W.W. (1959) Method for the Analysis of multicomponent exponential decay curves. J. Chem. Phys. 31: 978.

Garfinkel D. (1969) Simulation of Glycolytic Systems. In Heinmets F. (ed.) Concepts and Models of Biomathematics Marcel Dekker.

Garfinkel D., Williamson J.R., Olson M.S. (1969) Simulation of the Krebs Cycle. Simulation 12: 43.

Garfinkel D., McLeod J., Pring M., Di Toro D. (1972) Application of Computer Simulation to Research in the Life Sciences. Simulation 19: 17.

Gear C.W. (1971) The automatic integration of ordinary differential equations. Comm. A.C.M. 14: 176.

Geisler C.D., Hubbard A.E. (1972) New Boundary Conditions and Results for the Peterson-Bogert Model of the Cochlea. J. Acoustical Soc. Amer. 52: 1629.

Gelinas R.J. (1972) Stiff systems of kinetic equations - a practitioner's view. J. Comput. Phys. 9: 222.

Gibbons R.A., Sansom B.F., Sellwood R. (1972) The passage of calcium and strontium across the gut of the unaesthetized goat. J. Physiol. 222: 397.

Glass G.B.J. (1968) Introduction to Gastrointestinal Physiology, Prentice-Hall.

Glass H.I., de Garreta A.C. (1967) Quantitative analysis of exponential curve fitting for biological applications. Phys. Med. Biol. 12: 379.

Gleser M.A., Collen M.F. (1972) Towards Automated Medical Decisions. Cptrs and Biomed. Res. 5: 180.

Golomb S.W. (1971) Mathematical Models: Uses and Limitations. IEEE Trans. on Reliability R-20: 130.

Gonick H.C., Brown M. (1970) Critique of Multicompartmental analysis of calcium kinetics in man based on study of 27 cases. Metabolism: Clinical and Experimental 19: 919.

Goodman J.W. (1968) Introduction to Fourier Optics, McGraw-Hill.

Greville T.N.E. (1967) Spline functions, interpolation and numerical quadrature. In Ralston A., Wilf H.S. (eds) Mathematical Methods for Digital Computers vol. 2, Wiley.

Groner G.F., Clark R.L., Berman R.A., De Land E.C. (1971) Biomod: an interactive computer graphics system for modeling. Proceedings of the Fall Joint Computer Conference p.369.

Harmon L.D., Lewis E.R. (1968) Neural Modeling, in Levine S.N. (ed.) Advances in Biomedical Engineering and Medical Physics vol. 1, John Wiley.

Hart H., Spencer H. (1961) Rate of Initial Entry of Ca^{47} and Sr^{85} from the Intestine into the Vascular Space. Proc. Soc. Exp. Biol. Med. 126: 367.

Heaney R.P. (1963) Evaluation and Interpretation of Calcium Kinetic Data in Man. Clin. Orthop. 31: 153.

Heaney R.P., Skillman T.G. (1964) Secretion and excretion of calcium by the human gastrointestinal tract. J. Lab. and Clin. Med. 64: 29.

Hearon J.Z. (1963) Theorems on Linear Systems. Annals N.Y. Acad. Sci. no.108, p.36.

Helstrom C.W. (1967) Image restoration by the method of least squares. J. Opt. Soc. Amer. 57: 297.

Henrici P. (1962) Discrete variable methods in ordinary differential equations. Wiley.

Herscovitch H. and Schneider T. (1965) GPSS-III - An Expanded General Purpose Simulator. IBM Systems Journal 4: 174.

Hli F.B. (1971) A Time-Domain Approach. In Kalman R.E., De Claris N. (eds) Aspects of Network and System Theory, Holt, Rinehart and Winston.

Hopkins A. (1966) The pattern of gastric emptying: a new view of old results J. Physiol. 182: 144.

Hossfeld F. (1968) The correction of resolution errors in small angle scattering using Hermite functions. Acta Cryst. A24: 643.

Howard J.C., Young D.R. (1970) A simplified model of serum growth hormone regulation in man. Cptrs and Biomed. Res. 3: 101. Computers and Biomedical Research 3: 101.

Hsu F.T., Fan L.T., Hwang C.L. (1972) Simulation of a Steady-State Integrated Human Thermal System. Cptrs in Med. and Biol. 2: 59.

Hull T.E. (1969) The numerical integration of ordinary differential equations. IFIPS 1968 Conference, p.40.

Hunt J.N. (1973) Guy's Hospital Medical School London. Personal communication.

Hunt J.N. (1959) Gastric emptying and secretion in man. Physiol. Rev. 39: 491.

Hunt J.N., Knox M.T. (1967) Regulation of Gastric Emptying In Field J. (ed. in chief) Handbook of Physiology Section 6, Code C.F. (ed.) Alimentary Canal vol. 4 p.1917.

Hunt J.N., MacDonald I. (1954) The influence of volume on gastric emptying. J. Physiol. 126: 459.

I.B.M. (1968) System/360 Scientific Subroutine Package (360A-CM-03X) Version III Programmer's Manual. IBM Technical Publications Dept (NY).

Ibbertson H.K., Roche A.H.G., Pybus J. (1967) The thyroid and calcium homeostasis in Man: Evaluation by Calcium Infusion. Australasian Annals of Medicine 16: 121.

IEEE (1972) Special issue on digital picture processing. Proc. IEEE 60: 766.

Jackson A.S. (1960) Analogue Computation, McGraw-Hill.

Janes F.R., Carson E.R. (1971) Modelling of Biological Systems. Electronics and Power 17: 110.

Jansson P.A. (1970) Method for determining the response function of a high-resolution infrared spectrometer. J. Opt. Soc. Amer. 60: 184.

Jaworski Z.F., Brown E.M., Fedoruk S., Seitz H. (1963) A method for the study of calcium absorption by the human gut using a standard dose of calcium labelled with calcium^{47}. New England J. Med. 269: 1103.

Jenkins M.A., Traub J.F. (1972) Zeros of a complex polynomial Comm. A.C.M. 15: 97.

Jenkins M.A., Traub J.F. (1970) A three stage variable-shift iteration for polynomial zeros ans its relation to generalized Rayleigh iteration. Numerische Mathematik 14: 252.

Jones A.F., Misell D.L. (1970) The Problem of Error in Deconvolution. J. Phys. A: Gen. Phys. 3: 462.

Jones K.H., Fourman P. (1963) Edetic-acid test of parathyroid insufficiency. Lancet ii: 119.

Jones N.B., Porter N.H., Wood R.A. (1971) Modelling the Neuromuscular activity of the pelvic floor. In Conference on Computers for analysis and control in medical and biological research. Sept. 7-9 1971. IEE Conf. Publ. 79.

Jordan R.B. (1973) Computer Modelling of Adrenal Function. Ph.D. thesis, University of Canterbury, Christchurch, New Zealand, in preparation.

Kalman R.E. (1968) New Developments in Systems Theory Relevant to Biology. In Mesarovic M.D. (ed.) Systems Theory in Biology. Proc. of the III Systems Symposium at Case Institute of Technology. Springer-Verlag.

Kalman R.E. (1970) The Problem Setting in the Context of Linear-System Theory. In Kalman R.E. and de Claris N. (eds.) Aspects of Network and Systems Theory, Holt, Rinehart and Winston.

Kelly L.G. (1967) Handbook of Numerical Methods and Applications. Addison-Wesley.

Kivell J.A. (1973) The Analysis of X-Ray Diffraction Data Obtained from Deformed Metals with Special Reference to the Fatigue Behaviour of an Aluminium Alloy. M.E. Thesis University of Canterbury, Christchurch, New Zealand.

Kiviat P.J., Villanueva R., Markowitz H.M. (1968) The SIMSCRIPT II Programming Language. Prentice-Hall.

Klug H.P., Alexander L.E. (1967) X-Ray Diffraction Procedures, John Wiley.

Lanczos C. (1956) Applied Analysis, Prentice Hall.

Landahl H.D. (1972) Note on Oscillations in a Closed Chain of Compartments with Linear, Unidirectional Transport.

Bull. Math. Biophys. 34: 293.

Lathi B.P. (1967) Signals, Systems and Communication, John Wiley.

Lee R.C.K. (1964) Optimal Estimation, Identification and Control, Research Monograph no. 28 M.I.T. Press.

Lemaitre A., Malengé J-P. (1971) An Efficient Method for Multiexponential Fitting with a Computer. Cptrs. and Biomed. Res. 4: 555.

Lewis H.R., Stovall E.J. (1967) Comments on a floating point version of Nordsieck's scheme for the numerical integration of differential equations. Math. Comp. 21: 157.

Li Ching-Chung, Urquhart J. (1969) Modeling of Adrenocortical Secretory Dynamics. In Heinmets F. (ed.) Concepts and Models of Biomathematics, Marcel Dekker.

Liebowitz D., Stone H.H., Le Vine D., Scott K.G., Althausen T.L. (1957). Radioactive dilution indicator. I. Measurement of residual fluid in the fasting stomach. Gastroenterology 32: 265.

Livesey J.H. (1970) A Simulation Model of Calcium Metabolism. Ph.D. Thesis, University of Canterbury, Christchurch, New Zealand.

Louër P.D., Weigel D. (1969) Méthode Directe de Correction des Profils de Raies de Diffraction des Rayons X. I. Méthode Numerique de Deconvolution. Acta Cryst. A25: 335.

Lucas C.W., Terrill C.W. (1971) Algorithm 404 Complex Gamma Function [S14], Comm. A.C.M. 14: 48.

MacFadyen I.J., Nordin B.E.C., Smith D.A., Wayne D.J., Rae S.L. (1965) Effect of variation in dietary calcium on plasma concentration and urinary excretion of calcium. Brit. Med. J. i: 161.

MacFarlane A.G.J. (1964) Engineering Systems Analysis, George Harrap.

McLeod J. (1970) The Simulation of Difficult Systems. Simulation 14: 173.

Mains R.E., Soechting J.F. (1971) A Model for Neuromuscular Response to Sudden Disturbances. Trans. ASME Series G: 247.

Malm O.J. (1958) Calcium requirement and adaption in adult men. Scand. J. Clin. Lab. Invest. 10 (suppl. 36): 1.

Mancini P., Pilo A. (1970) A computer program for multi-exponential fitting by the peeling method. Cptrs and Biomed. Res. 3: 1.

Mandel J. (1964) The Statistical Analysis of Experimental Data. John Wiley.

Marshall D.H., Nordin B.E.C. (1969) Kinetic analysis of plasma radioactivity after oral injestion of radio calcium. Nature 222: 797.

Marshall J.H. (1964) Theory of alkaline earth metabolism. J. Theoret. Biol. 6: 386.

Marshall J.H., Onkelinx C. (1968) Radial diffusion and power function retention of alkaline earth radioisotopes in bone. Nature 217: 742.

Massin J.-P., Savoie J.-C., Camus J.-P., Lièvre J.-A. Étude isotopique du métabolisme calcique chez 5 sujets normaux et 5 sujets atteints d'hyperparathyroïdisme primitif Essai d'une méthode d'analyse tricompartimentale. Annals d'Endocrinologie 29: 211.

Mautalen C.A., Cabrejas M.L., Soto R.J. (1969) Isotopic determination of intestinal calcium absorption in normal subjects. Metab.: Clin. and Expt. 18: 395.

Mesarovic M.D. (1968) Systems Theory and Biology - View of a Theoretician. In Mesarovic M.D. (ed.) Systems Theory and Biology. Springer-Verlag.

Mesarovic M.D., Macko D. Takahara Y. (1970) The Theory of Hierarchical, Multilevel Systems. Academic Press.

Metcalf W.S. (1973) Chemistry Dept, University of Canterbury, Christchurch, New Zealand. Private communication.

Mihram G.A. (1972) The Modelling Process. IEEE Trans. on Systems, Man and Cybernetics SMC-2: 621.

Milsum J.H. (1970) Biological Systems Analysis and Control Theory. In Clynes M. and Milsum J.H. (eds) Biomedical Engineering Systems. McGraw-Hill.

Moore L. (1968) Deconvolution of Physical Data. Brit. J. Appl. Phys. (J. Phys. D) ser. 2,1: 237.

Morse P.M., Feshbach H. (1953). Methods of Theoretical Physics McGraw-Hill.

Myhill J. (1968) Some effects of data error in the analysis of radiotracer data. Acta Radiol. Ther. Phys. Biol. 7: 443.

Myhill J., Wadsworth G.P., Brownell G.L. (1965) Investigation of an operator method in the analysis of biological tracer data. Biophys. J. 5: 89.

Napier P.J. (1972) Reconstruction of Radiating Sources. Ph.D. Thesis, University of Canterbury, Christchurch, New Zealand.

Neer R., Berman M., Fisher L., Rosenberg L.E. (1967) Multi-compartmental analysis of calcium kinetics in normal adult males. J. Clin. Invest. 46: 1364-79.

Neufeld G.R. (1971) Computation of transit time distributions using sampled data Laplace transforms. J. Appl. Physiol. 31: 148.

Nieman R.E., Fisher D.G., Seborg D.E. (1971) A review of process identification and parameter estimation techniques. Int. J. Control 13: 209.

Nordin B.E.C., Fraser R. (1954) The Effect of Intravenous Calcium on Phosphate Excretion. Clin. Sci. 13: 477.

Nordin B.E.C., Smith D.A. (1965) Diagnostic Procedures in Disorders of Calcium Metabolism. Churchill.

Nordsieck A. (1962) On Numerical Integration of Ordinary Differential Equations. Comp. Math. 16: 22.

O'Brien M.M., McIntosh H.W. (1967) Observations relating to the possible role of calcitonin in calcium homeostasis in man. Can. Med. Assn J. 97: 941.

Paley R.E.A.C., Wiener N. (1934) Fourier transforms in the Complex Domain. American Math. Soc.

Parsons D.H. (1968) Biological Problems involving sums of exponential functions of time: A mathematical analysis that reduces experimental time. Math. Biosci. 2: 123.

Partridge L.D. (1972) Simulation in Biomedical Teaching. IEEE Trans. on Biomed. Eng. BME-19: 78.

Patterson M.S. (1950) Calculation of the correction for instrumental broadening in X-ray diffraction lines. Proc. Phys. Soc. A63: 477.

Pearson A.J. (1972) The Dynamics of Calcium Transfer in Man M.Sc. thesis, University of Canterbury, Christchurch, New Zealand.

Perl W. (1960) A method for curve-fitting by exponential functions. Int. J. Appl. Radiat. Isotopes 8: 211.

Phang J.M., Berman M., Finerman G.A., Neer R.M., Rosenberg L.E., Hahn T. (1969) Dietary perturbation of calcium metabolism in normal man: Compartmental analysis. J. Clin. Invest. 48: 67.

Piessens R. (1971) On a numerical method for the calculation of transient responses. J. of the Franklin Institute 292: 57.

Piessens R., Collen D., Tytgat G. (1971) Computer Analysis of Fibrogen Tracer Data. Cptrs and Biomed. Res. 4: 585.

Pizer S.M., Ashare A.B., Callahan A.B., Brownell G.L. (1969) Fourier Transform Analysis of Tracer Data. In Heinmets F. (ed.) Concepts and Models of Biomathematics p.105. Marcel Dekker.

Powell T. (1972) A Mathematical Model for Calcium Homeostasis. Bull. Math. Biophys. 34: 483.

Pritsker A.A.B., Kiviat P.J. (1969) Simulation with Gasp II. Prentice-Hall.

Ramberg C.F., Mayer G.P., Kronfeld D.S., Phang J.M., Berman M. (1970) Calcium kinetics in cows during late pregnancy, parturition and early lactation. Amer. J. Physiol. 219: 1166.

Renken J.H., Biggs F. (1972) Use of numerically inverted Laplace transforms in time-dependent transport calculations. J. Comput. Phys. 9: 318.

Rescigno A. Segre G. (1962) Analysis of multicompartmented biological systems. J. Theoret. Biol. 3: 149.

Riggs D.S. (1966) A quantitative hypothesis concerning the action of the parathyroid hormone. J. Theoret. Biol. 12: 364.

Robertson-Dunn B., Ng K.K., Kwang K., Linkens D.A. (1971) Computer modelling of electrical activity of the gastro-intestinal tract. in conference on Computers for Analysis and Control in Medical and Biological Research, 7-9 Sept. 1971. IEE Conf. Publ. No. 79.

Roman P., Marathay A.S. (1963) Analyticity and Phase Retrieval. Nuovo Cimento 30: 6010.

Roston S. (1959) Mathematical representation of some endocrinological systems. Bull. Math. Biophys. 21: 271.

Schoeffler J.D., Ostrander L.E., Gann D.S. (1968) Identification of Boolean Mathematical Models. In Mesarović M.D. (ed.) Systems Theory and Biology, Springer-Verlag.

Segre G., Turco G.L., Vercellone G. (1973) Modeling Blood Glucose and Insulin Kinetics in Normal Diabetic and Obese Subjects. Diabetes 22: 94.

Sharney L., Tendler D., Stevenson J.C., Wasserman L.R. (1971) Multiple-pool analysis of metabolic pathways. Med. Res. Eng. 10: 4.

Sheiner L.B., Rosenberg B. Melmon K.L. (1972) Modelling of Individual Pharmacokinetics for Computer-Aided Drug Dosage. Cptrs and Biomed. Res. 5: 441.

Sheppard C.W. (1948) The theory of the study of transfers within a multicompartment system using isotopic tracers. J. Appl. Phys. 19: 70.

Shipley R.A., Clark R.E. (1972) Tracer Methods for In Vivo Kinetics Theory and Applications, Academic Press.

Silverberg M. (1970) Improving the efficiency of Laplace-transform inversion for network analysis. Electronics Letters 6: 105.

Silverman M., Burgen A.S.V. (1961) Application of the analogue computer to measurement of intestinal absorption rates with tracers. J. Appl. Physiol. 16: 911.

Silverman H.F., Pearson A.E. (1973) On deconvolution using the discrete Fourier transform. IEEE Trans. on Automatic Control AU-21: 112.

Sondhi M.M. (1972) Image Restoration: The removal of spatially invariant degradation. Proc. IEEE 60: 842.

Spyker D.A. (1970) Simulation in the Analysis and Control of a cardio-circulatory assist device. Simulation 15: 196.

Stokely E.M., Howard L.L. (1972) Analogue computer model for the ACTH glucocorticoid system. IEEE Trans. on Biomedical Engineering BME-19: 13.

Stokes A.R. (1948) A numerical Fourier-analysis method for the correction of widths and shapes of lines on X-ray powder photographs. Proc. Phys. Soc. (London) 61: 382.

Szymendera J., Heaney R.P., Saville P.D. (1972) Intestinal Calcium Absorption: Concurrent use of oral and intravenous tracers and calculation by the inverse convolution method. J. Lab. Clin. Med. 79: 570.

Taha H.A. (1971) Operations Research. An Introduction. MacMillan.

Thron C.D. (1972) Structure and kinetic behaviour of linear multicompartment systems. Bull. Math. Biophys. 34: 277.

Tomović R. (1963) Sensitivity Analysis of Dynamic Systems. McGraw-Hill.

Treanor C.E. (1966) A method for the numerical integration of coupled first-order differential equations with greatly different time constants. Math. Comp. 20: 39.

Tuttle D.F. (1971) On Fluids, Networks and Engineering Education. In Kalman R.E. and De Claris N. (eds) Aspects of Network and Systems Theory. Holt, Rinehart and Winston.

Vaughan J.M. (1970) The Physiology of Bone. Clarendon Press.

Wagner J.G. (1968) Pharmacokinetics. Annual Review of Pharmacology 8: 67-94.

Warner H.R., Rutherford B.D., Houtchens B. (1972) A Sequential Bayesian Approach to History Taking and Diagnosis. Cptrs and Biomed. Res. 5: 256.

Waterman T.H. (1968) Systems Theory and Biology - View of a Biologist. In Mesarović M.D. (ed.) Systems Theory and Biology. Springer-Verlag.

Watt K. and Staff (1970) A model of society. Simulation 14: 153.

Wensel R.H., Rich C., Brown A.C., Wolwiler W. (1969) Absorption of calcium measured by intubation and perfusion of the intact human small intestine. J. Clin. Invest. 48: 1768.

Westlake J.R. (1968) A Handbook of Numerical Matrix Inversion and Solution of Linear Equations. John Wiley.

Whalen G.E., Harris J.A., Greenen J.E., Soergel K.H. (1966) Sodium and water absorption from the human small intestine. Gastroenterology 51: 975.

Wigan M.R. (1972) The fitting, calibration and validation of simulation models. Simulation 18: 188.

Wilson R. (1970) A study of the methods of parameter estimation for multicompartment models of physiological and biological systems. M.Sc. Thesis. Dept of Control and Automation, City University, London.

Wolaver L.E. (1972) Mathematical Model for Calcium Homeostasis. 5th International Conference on System Sciences Hawaii 1972. Proceedings Supplement, p.153.

Zadeh L.A. (1970) Toward a Theory of Fuzzy Systems. In Kalman R.E., de Claris N. (eds.) Aspects of Network and System Theory. Holt, Rinehart and Winston.

Zakian V. (1970) Optimisation of Numerical Inversion of Laplace Transforms. Electronics Letters 6: 678.